# PATENT COOPERATION TREATY

# PCT



## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

| Applicant's or agent's file reference | FOR FURTHER ACTION | See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416) |
|---|---|---|
| BB1159 | | |
| International application No. | International filing date *(day/month/year)* | Priority date *(day/month/year)* |
| PCT/US99/16354 | 20/07/1999 | 21/07/1998 |
| International Patent Classification (IPC) or national classification and IPC<br>C12N15/60 | | |
| Applicant<br>E.I. DU PONT DE NEMOURS AND COMPANY et al. | | |

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☒ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

| Date of submission of the demand | Date of completion of this report |
|---|---|
| 07/02/2000 | 05.09.2000 |
| Name and mailing address of the international preliminary examining authority:<br>European Patent Office<br>D-80298 Munich<br>Tel. +49 89 2399 - 0 Tx: 523656 epmu d<br>Fax: +49 89 2399 - 4465 | Authorized officer<br>Schwachtgen, J-L<br>Telephone No. +49 89 2399 8933 |

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/16354

## I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

**Description, pages:**

1-23 as originally filed

**Claims, No.:**

1-12 as originally filed

**Drawings, sheets:**

1/3-3/3 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

## IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

☐ restricted the claims.

☐ paid additional fees.

☐ paid additional fees under protest.

☐ neither restricted nor paid additional fees.

2. ☒ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

   ☐ complied with.

   ☒ not complied with for the following reasons:

   **see separate sheet**

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

   ☒ all parts.

   ☐ the parts relating to claims Nos. .

## V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Yes: | Claims | 1-5, 7-9, 12 |
| | No: | Claims | 6, 10, 11 |
| Inventive step (IS) | Yes: | Claims | |
| | No: | Claims | 1-12 |
| Industrial applicability (IA) | Yes: | Claims | 1-12 |
| | No: | Claims | |

2. Citations and explanations

   **see separate sheet**

**Re Item IV**

**Lack of unity of invention**

The present application does not meet the requirements for unity of invention (Rule 13.1, 13.2 and 13.3 PCT):

The present claims have been divided into 5 groups of invention (PCT ISA 210 dated 08/12/1999).

2. Rule 13 PCT states that for unity of invention to be present all subject-matter should be linked by a single general concept. Rule 13 PCT also stipulates that this single general concept must be novel and inventive.

2.1 In the present case the only linking concept that can be identified is that all groups address a problem which can be formulated as "providing isolated nucleic acid fragments encoding 3-DEHYDROQUINATE SYNTHASE".

This common problem (concept), was already known from D1.

2.1 Furthermore, the documents cited in the Search Report, taken jointly and severally, illustrate that provision of 3-DEHYDROQUINATE SYNTHASE was well known at the priority date.

2.2 The linking concept is therefore neither novel nor inventive and cannot provide unity of invention for the present Groups of invention.

**Re Item V**

**Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

Reference is made to the following documents:

D1: MILLAR G ET AL: 'THE COMPLETE AMINO ACID SEQUENCE OF 3-

DEHYDROQUINATE SYNTHASE OF ESCHERICHIA COLI K12' FEBS LETTERS, vol. 200, no. 1, page 11-17 XP002910008 ISSN: 0014-5793 & MILLAR, G., ET AL.: '3-dehydroquinate synthase(E.C4.6.1.3)' SWISSPROT ACCESSION NO:P07639, 1 April 1998 (1998-04-01)

D2: BISCHOFF, M., ET AL.: 'Cloning of a cDNA encoding a 3-dehydroquinate synthase gene from higher plant, and analyis of the organ-specific and elicitor-induced expression of the corresponding gene' PLANT MOLECULAR BIOLOGY, vol. 31, 1996, pages 69-76, XP002123455

1. The present application does not meet the requirements set forth in Article 33(2) PCT because the subject-matter of claims 10 and 11 is not new in respect of the prior art as defined in the regulations (Rule 64(1)-(3) PCT).

D1 discloses the isolated nucleic acid encoding a 3-dehydroquinate synthase (DHQ) from E-coli (abstract). Said enzyme has extensive regions of identity with the DHQ sequences claimed in the present application. It is obtainable by the screening method according to claims 8 and 9 and, Thus, it anticipates the novelty of the subject-matter of claims 10 and 11.
The same objection applies to the subject-matter of claim 6.

2. The subject-matter of claim 1 relates to DHQ enzymes which have not been disclosed in the prior art.

The document D2, which is regarded as being the closest prior art to the subject-matter of claim 1, discloses a cDNA encoding a 3-dehydroquinate synthase from tomato.

The technical problem may therefore be regarded as the provision of further isolated nucleic acid molecules encoding 3-dehydroquinate synthase.

The solution proposed in claim 1 of the present application cannot be considered

as involving an inventive step (Article 33(3) PCT) because the chosen sequences are not shown to possess any common advantageous properties not possessed by the prior art examples (see PCT Guidelines, C-IV, 8.8, C1, iv).

The same objection applies to the subject-matter of claims 2-12 as the role of DHQ as a target for herbicides was known from D3.

# PATENT COOPERATION TREATY

SEP 21 2000

RECEIVED

SEP 13 2000

PATENT RECORDS CENTER

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

LYNNE M. Christenbury
E.I. Du Pont de Nemours and Company
Legal Patent Records Center
1007 Market Street
Wilmington, DE 19898
ETATS-UNIS D'AMERIQUE

KL

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT
(PCT Rule 71.1)

| | |
|---|---|
| Date of mailing *(day/month/year)* | 05.09.2000 |
| Applicant's or agent's file reference: BB1159 | **IMPORTANT NOTIFICATION** |

| International application No. | International filing date *(day/month/year)* | Priority date *(day/month/year)* |
|---|---|---|
| PCT/US99/16354 | 20/07/1999 | 21/07/1998 |

Applicant
E.I. DU PONT DE NEMOURS AND COMPANY et al.

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

TRB NOTED

| Name and mailing address of the IPEA/ | Authorized officer |
|---|---|
| European Patent Office<br>D-80298 Munich<br>Tel. +49 89 2399 - 0 Tx: 523656 epmu d<br>Fax: +49 89 2399 - 4465 | Vullo, C<br><br>Tel.+49 89 2399-8061 |

Form PCT/IPEA/416 (July 1992)

21 JAN 2001

# PATENT COOPERATION TREATY

# PCT

## INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

| Applicant's or agent's file reference | FOR FURTHER ACTION | See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416) |
|---|---|---|
| BB1159 | | |
| International application No. | International filing date *(day/month/year)* | Priority date *(day/month/year)* |
| PCT/US99/16354 | 20/07/1999 | 21/07/1998 |
| International Patent Classification (IPC) or national classification and IPC<br>C12N15/60 | | |
| Applicant<br>E.I. DU PONT DE NEMOURS AND COMPANY et al. | | |

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   ☐ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of sheets.

3. This report contains indications relating to the following items:

   - I ☒ Basis of the report
   - II ☐ Priority
   - III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   - IV ☒ Lack of unity of invention
   - V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   - VI ☐ Certain documents cited
   - VII ☐ Certain defects in the international application
   - VIII ☐ Certain observations on the international application

| Date of submission of the demand | Date of completion of this report |
|---|---|
| 07/02/2000 | 05.09.2000 |
| Name and mailing address of the international preliminary examining authority:<br>European Patent Office<br>D-80298 Munich<br>Tel. +49 89 2399 - 0 Tx: 523656 epmu d<br>Fax: +49 89 2399 - 4465 | Authorized officer<br>Schwachtgen, J-L<br>Telephone No. +49 89 2399 8933 |

Form PCT/IPEA/409 (cover sheet) (January 1994)

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/US99/16354

## I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

**Description, pages:**

1-23 as originally filed

**Claims, No.:**

1-12 as originally filed

**Drawings, sheets:**

1/3-3/3 as originally filed

2. The amendments have resulted in the cancellation of:

☐ the description, pages:
☐ the claims, Nos.:
☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

## IV. Lack of unity of invention

1. In response to the invitation to restrict or pay additional fees the applicant has:

☐ restricted the claims.

☐ paid additional fees.

☐ paid additional fees under protest.

☐ neither restricted nor paid additional fees.

2. ☒ This Authority found that the requirement of unity of invention is not complied and chose, according to Rule 68.1, not to invite the applicant to restrict or pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rules 13.1, 13.2 and 13.3 is

   ☐ complied with.

   ☒ not complied with for the following reasons:

   **see separate sheet**

4. Consequently, the following parts of the international application were the subject of international preliminary examination in establishing this report:

   ☒ all parts.

   ☐ the parts relating to claims Nos. .

## V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Yes: | Claims | 1-5, 7-9, 12 |
| | No: | Claims | 6, 10, 11 |
| Inventive step (IS) | Yes: | Claims | |
| | No: | Claims | 1-12 |
| Industrial applicability (IA) | Yes: | Claims | 1-12 |
| | No: | Claims | |

2. Citations and explanations

   **see separate sheet**

**Re Item IV**

**Lack of unity of invention**

The present application does not meet the requirements for unity of invention (Rule 13.1, 13.2 and 13.3 PCT):

The present claims have been divided into 5 groups of invention (PCT ISA 210 dated 08/12/1999).

2. Rule 13 PCT states that for unity of invention to be present all subject-matter should be linked by a single general concept. Rule 13 PCT also stipulates that this single general concept must be novel and inventive.

2.1 In the present case the only linking concept that can be identified is that all groups address a problem which can be formulated as "providing isolated nucleic acid fragments encoding 3-DEHYDROQUINATE SYNTHASE".

This common problem (concept), was already known from D1.

2.1 Furthermore, the documents cited in the Search Report, taken jointly and severally, illustrate that provision of 3-DEHYDROQUINATE SYNTHASE was well known at the priority date.

2.2 The linking concept is therefore neither novel nor inventive and cannot provide unity of invention for the present Groups of invention.

**Re Item V**

**Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

Reference is made to the following documents:

D1: MILLAR G ET AL: 'THE COMPLETE AMINO ACID SEQUENCE OF 3-

DEHYDROQUINATE SYNTHASE OF ESCHERICHIA COLI K12' FEBS LETTERS, vol. 200, no. 1, page 11-17 XP002910008 ISSN: 0014-5793 & MILLAR, G., ET AL.: '3-dehydroquinate synthase(E.C4.6.1.3)' SWISSPROT ACCESSION NO:P07639, 1 April 1998 (1998-04-01)

D2: BISCHOFF, M., ET AL.: 'Cloning of a cDNA encoding a 3-dehydroquinate synthase gene from higher plant, and analyis of the organ-specific and elicitor-induced expression of the corresponding gene' PLANT MOLECULAR BIOLOGY, vol. 31, 1996, pages 69-76, XP002123455

1. The present application does not meet the requirements set forth in Article 33(2) PCT because the subject-matter of claims 10 and 11 is not new in respect of the prior art as defined in the regulations (Rule 64(1)-(3) PCT).

   D1 discloses the isolated nucleic acid encoding a 3-dehydroquinate synthase (DHQ) from E-coli (abstract). Said enzyme has extensive regions of identity with the DHQ sequences claimed in the present application. It is obtainable by the screening method according to claims 8 and 9 and, Thus, it anticipates the novelty of the subject-matter of claims 10 and 11.
   The same objection applies to the subject-matter of claim 6.

2. The subject-matter of claim 1 relates to DHQ enzymes which have not been disclosed in the prior art.

   The document D2, which is regarded as being the closest prior art to the subject-matter of claim 1, discloses a cDNA encoding a 3-dehydroquinate synthase from tomato.

   The technical problem may therefore be regarded as the provision of further isolated nucleic acid molecules encoding 3-dehydroquinate synthase.

   The solution proposed in claim 1 of the present application cannot be considered

as involving an inventive step (Article 33(3) PCT) because the chosen sequences are not shown to possess any common advantageous properties not possessed by the prior art examples (see PCT Guidelines, C-IV, 8.8, C1, iv).

The same objection applies to the subject-matter of claims 2-12 as the role of DHQ as a target for herbicides was known from D3.

# PATENT COOPERATION TREATY

From the INTERNATIONAL SEARCHING AUTHORITY

# PCT

To:
E.I. DU PONT DE NEMOURS AND COMPANY
Legal/Patent Records Center
Attn. MAJARIAN, William R.
1007 Market Street
Wilmington, Delaware 19898
UNITED STATES OF AMERICA



NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL SEARCH REPORT OR THE DECLARATION

(PCT Rule 44.1)

| | |
|---|---|
| Applicant's or agent's file reference<br>BB1159A | Date of mailing *(day/month/year)*<br>10/03/2000 |
| International application No.<br>PCT/US 99/ 16353 | **FOR FURTHER ACTION** See paragraphs 1 and 4 below |
| Applicant<br>E.I. DU PONT DE NEMOURS AND COMPANY et al. | International filing date *(day/month/year)*<br>20/07/1999 |

1. [X] The applicant is hereby notified that the International Search Report has been established and is transmitted herewith.

**Filing of amendments and statement under Article 19:**
The applicant is entitled, if he so wishes, to amend the claims of the International Application (see Rule 46):

**When?** The time limit for filing such amendments is normally 2 months from the date of transmittal of the International Search Report; however, for more details, see the notes on the accompanying sheet.

**Where?** Directly to the International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland
Fascimile No.: (41-22) 740.14.35

**For more detailed instructions,** see the notes on the accompanying sheet.

2. [ ] The applicant is hereby notified that no International Search Report will be established and that the declaration under Article 17(2)(a) to that effect is transmitted herewith.

3. [ ] **With regard to the protest** against payment of (an) additional fee(s) under Rule 40.2, the applicant is notified that:

[ ] the protest together with the decision thereon has been transmitted to the International Bureau together with the applicant's request to forward the texts of both the protest and the decision thereon to the designated Offices.

[ ] no decision has been made yet on the protest; the applicant will be notified as soon as a decision is made.

4. **Further action(s):** The applicant is reminded of the following:

Shortly after **18 months** from the priority date, the international application will be published by the International Bureau. If the applicant wishes to avoid or postpone publication, a notice of withdrawal of the international application, or of the priority claim, must reach the International Bureau as provided in Rules 90*bis*.1 and 90*bis*.3, respectively, before the completion of the technical preparations for international publication.

Within **19 months** from the priority date, a demand for international preliminary examination must be filed if the applicant wishes to postpone the entry into the national phase until 30 months from the priority date (in some Offices even later).

Within **20 months** from the priority date, the applicant must perform the prescribed acts for entry into the national phase before all designated Offices which have not been elected in the demand or in a later election within 19 months from the priority date or could not be elected because they are not bound by Chapter II.

| | |
|---|---|
| Name and mailing address of the International Searching Authority<br>European Patent Office, P.B. 5818 Patentlaan 2<br>NL-2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Authorized officer<br>Mireille Claudepierre<br>CLS NOTED 3/16/2000 |

Form PCT/ISA/220 (July 1998)

# NOTES TO FORM PCT/ISA/220

These Notes are intended to give the basic instructions concerning the filing of amendments under article 19. The Notes are based on the requirements of the Patent Cooperation Treaty, the Regulations and the Administrative Instructions under that Treaty. In case of discrepancy between these Notes and those requirements, the latter are applicable. For more detailed information, see also the PCT Applicant's Guide, a publication of WIPO.

In these Notes, "Article", "Rule", and "Section" refer to the provisions of the PCT, the PCT Regulations and the PCT Administrative Instructions, respectively.

## INSTRUCTIONS CONCERNING AMENDMENTS UNDER ARTICLE 19

The applicant has, after having received the international search report, one opportunity to amend the claims of the international application. It should however be emphasized that, since all parts of the international application (claims, description and drawings) may be amended during the international preliminary examination procedure, there is usually no need to file amendments of the claims under Article 19 except where, e.g. the applicant wants the latter to be published for the purposes of provisional protection or has another reason for amending the claims before international publication. Furthermore, it should be emphasized that provisional protection is available in some States only.

**What parts of the international application may be amended?**

Under Article 19, only the claims may be amended.

During the international phase, the claims may also be amended (or further amended) under Article 34 before the International Preliminary Examining Authority. The description and drawings may only be amended under Article 34 before the International Examining Authority.

Upon entry into the national phase, all parts of the international application may be amended under Article 28 or, where applicable, Article 41.

**When?**

Within 2 months from the date of transmittal of the international search report or 16 months from the priority date, whichever time limit expires later. It should be noted, however, that the amendments will be considered as having been received on time if they are received by the International Bureau after the expiration of the applicable time limit but before the completion of the technical preparations for international publication (Rule 46.1).

**Where not to file the amendments?**

The amendments may only be filed with the International Bureau and not with the receiving Office or the International Searching Authority (Rule 46.2).

Where a demand for international preliminary examination has been/is filed, see below.

**How?**

Either by cancelling one or more entire claims, by adding one or more new claims or by amending the text of one or more of the claims as filed.

A replacement sheet must be submitted for each sheet of the claims which, on account of an amendment or amendments, differs from the sheet originally filed.

All the claims appearing on a replacement sheet must be numbered in Arabic numerals. Where a claim is cancelled, no renumbering of the other claims is required. In all cases where claims are renumbered, they must be renumbered consecutively (Administrative Instructions, Section 205(b)).

**The amendments must be made in the language in which the international application is to be published.**

**What documents must/may accompany the amendments?**

**Letter (Section 205(b)):**

The amendments must be submitted with a letter.

The letter will not be published with the international application and the amended claims. It should not be confused with the "Statement under Article 19(1)" (see below, under "Statement under Article 19(1)").

**The letter must be in English or French, at the choice of the applicant. However, if the language of the international application is English, the letter must be in English; if the language of the international application is French, the letter must be in French.**

## NOTES TO FORM PCT/ISA/220 (continue)

The letter must indicate the differences between the claims as filed and the claims as amended. It must, in particular, indicate, in connection with each claim appearing in the international application (it being understood that identical indications concerning several claims may be grouped),whether

(i) the claim is unchanged;

(ii) the claim is cancelled;

(iii) the claim is new;

(iv) the claim replaces one or more claims as filed;

(v) the claim is the result of the division of a claim as filed.

**The following examples illustrate the manner in which amendments must be explained in the accompanying letter:**

1. [Where originally there were 48 claims and after amendment of some claims there are 51]:
"Claims 1 to 29, 31, 32, 34, 35, 37 to 48 replaced by amended claims bearing the same numbers; claims 30, 33 and 36 unchanged; new claims 49 to 51 added."

2. [Where originally there were 15 claims and after amendment of all claims there are 11]:
"Claims 1 to 15 replaced by amended claims 1 to 11."

3. [Where originally there were 14 claims and the amendments consist in cancelling some claims and in adding new claims]:
"Claims 1 to 6 and 14 unchanged; claims 7 to 13 cancelled; new claims 15, 16 and 17 added." or
"Claims 7 to 13 cancelled; new claims 15, 16 and 17 added; all other claims unchanged."

4. [Where various kinds of amendments are made]:
"Claims 1-10 unchanged; claims 11 to 13, 18 and 19 cancelled; claims 14, 15 and 16 replaced by amended claim 14; claim 17 subdivided into amended claims 15, 16 and 17; new claims 20 and 21 added."

**"Statement under article 19(1)" (Rule 46.4)**

The amendments may be accompanied by a statement explaining the amendments and indicating any impact that such amendments might have on the description and the drawings (which cannot be amended under Article 19(1)).

The statement will be published with the international application and the amended claims.

**It must be in the language in which the international application is to be published.**

It must be brief, not exceeding 500 words if in English or if translated into English.

It should not be confused with and does not replace the letter indicating the differences between the claims as filed and as amended. It must be filed on a separate sheet and must be identified as such by a heading, preferably by using the words "Statement under Article 19(1)."

It may not contain any disparaging comments on the international search report or the relevance of citations contained in that report. Reference to citations, relevant to a given claim, contained in the international search report may be made only in connection with an amendment of that claim.

**Consequence if a demand for international preliminary examination has already been filed**

If, at the time of filing any amendments and any accompanying statement, under Article 19, a demand for international preliminary examination has already been submitted, the applicant must preferably, at the time of filing the amendments (and any statement ) with the International Bureau, also file with the International Preliminary Examining Authority a copy of such amendments (and of any statement) and, where required, a translation of such amendments for the procedure before that Authority (see Rules 55.3(a) and 62.2, first sentence). For further information, see the Notes to the demand form (PCT/IPEA/401).

**Consequence with regard to translation of the international application for entry into the national phase**

The applicant's attention is drawn to the fact that, upon entry into the national phase, a translation of the claims as amended under Article 19 may have to be furnished to the designated/elected Offices, instead of, or in addition to, the translation of the claims as filed.

For further details on the requirements of each designated/elected Office, see Volume II of the PCT Applicant's Guide.

# PATENT COOPERATION TREATY

# PCT

## INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

| Applicant's or agent's file reference | **FOR FURTHER ACTION** | see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below. |
|---|---|---|
| BB1159A | | |
| International application No. | International filing date *(day/month/year)* | (Earliest) Priority Date *(day/month/year)* |
| PCT/US 99/ 16353 | 20/07/1999 | 21/07/1998 |
| Applicant | | |
| E.I. DU PONT DE NEMOURS AND COMPANY et al. | | |

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 6 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.

[ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).

b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :

[X] contained in the international application in written form.

[X] filed together with the international application in computer readable form.

[ ] furnished subsequently to this Authority in written form.

[ ] furnished subsequently to this Authority in computer readble form.

[ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.

[ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).

3. [X] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

[X] the text is approved as submitted by the applicant.

[ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. -

[ ] as suggested by the applicant.

[ ] because the applicant failed to suggest a figure.

[ ] because this figure better characterizes the invention.

[ ] None of the figures.

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/16353

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☒ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

**Remark on Protest**

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-12 all partially

An isolated nucleic acid fragment encoding a chorismate synthase which hybridizes to a nucleic acid encoding an amino acid set forth in SEQ ID NO: 2 and 8, said nucleic acid consisiting of SEQ ID NO: 1 and 7, a chorismate synthase polypeptide comprising all or portion of SEQ ID NO: 2 and 8. A chimeric gene, a transformed host, a method for altering the level of expression of a chorismate synthase, methods for obtaining a nucleic acid fragment, products obtained by said methods, a method for evaluating compounds inhibiting chorismate synthase utilizing said nucleic acids and polypeptides.

2. Claims: 1-12 all partially

same as invention 1 but comprising SEQ ID NO: 3, 4, 13 and 14.

3. Claims: 1-12 all partially

same as invention 1 but comprising SEQ ID NO: 5 and 6.

4. Claims: 1-12 all partially

same as invention 1 but comprising SEQ ID NO: 9-12.

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/16353

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N9/88 C12Q1/68 C12Q1/527 C12N15/82

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | GÖRLACH J ET AL: "Differential expression of tomato (Lycopersicon esculentum L.) genes encoding shikimate pathway isoenzymes. II. Chorismate synthase." PLANT MOLECULAR BIOLOGY, vol. 23, no. 4, November 1993 (1993-11), pages 707-716, XP002120076 cited in the application the whole document --- -/-- | 1-11 |

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 25 February 2000 | 10.03.00 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,<br>Fax: (+31-70) 340-3016 | Oderwald, H |

2

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | SCHALLER A ET AL: "Molecular cloning and analysis of a cDNA coding for chorismate synthase from the higher plant Corydalis sempervirens Pers." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 266, no. 32, 15 November 1991 (1991-11-15), pages 21434-21438, XP002120077 the whole document | 1-11 |
| X | WO 98 03661 A (ROBERTS FIONA ;ROBERTS CRAIG W (GB); ARCH DEV CORP (US); MCLEOD RI) 29 January 1998 (1998-01-29) abstract; claims 1,2,4,5,8,12,13,26,27; figures 4,5,9,10; examples 1-8,12-19,25,33,38 page 1 -page 41 | 1-12 |
| X | DATABASE EMEST14 [Online] EMBL, Hinxton, GB AC/ID: C72774, 19 September 1997 (1997-09-19) YAMAMOTO K AND SASAKI T: "Rice cDNA from panicle at flowering stage (970813)" XP002120079 abstract | 1-5,7-11 |
| X | DATABASE EMEST3 [Online] EMBL, Heidelberg, Germany AC/ID AA750226, 21 January 1998 (1998-01-21) NAHM B H ET AL: "Large-scale sequencing analysis of ESTs from rice immature seed" XP002131647 abstract | 1-5,7-11 |
| X | BORNEMANN S ET AL.: "Escherichia coli chorismate synthase catalyzes the conversion of (6S)-6-fluoro-5-enolpyruvylshikimate-3-phosphate to 6-fluorochorismate" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 270, no. 39, 29 September 1995 (1995-09-29), pages 22811-22815, XP002120078 the whole document | 12 |
| P,X | DATABASE EMEST12 [Online] EMBL, Heidelberg, Germany AC/ID AI731017, 12 June 1999 (1999-06-12) BLEWITT M ET AL.: "ESTs from developing cotton fiber" XP002131648 abstract | 1-5,7-11 |

-/--

2

# INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/16353

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category ° | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | DATABASE EMEST12 [Online]<br>EMBL, Heidelberg, Germany<br>AC/ID AI728073, 12 June 1999 (1999-06-12)<br>BLEWITT M ET AL.: "ESTs from developing cotton fiber"<br>XP002131649<br>abstract<br>---<br>DATABASE EMEST10 [Online]<br>AC/ID AI489566, 17 March 1999 (1999-03-17)<br>ALCALA J ET AL.: "Generation of ESTs from tomato carpel tissue"<br>XP002131650<br>abstract<br>----- | 1-5,7-11<br><br>1-5,7-11 |

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/16353

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| WO 9803661 A | 29-01-1998 | AU 4041197 A | 10-02-1998 |
| | | EP 0918868 A | 02-06-1999 |

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/60, 1/21, C12Q 1/68, 1/527

A1

(11) International Publication Number: WO 00/05387

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/US99/16354

(22) International Filing Date: 20 July 1999 (20.07.99)

(30) Priority Data:
60/093,611 21 July 1998 (21.07.98) US

(71) Applicant *(for all designated States except US)*: E.I. DU PONT DE NEMOURS AND COMPANY [US/US]; 1007 Market Street, Wilmington, DE 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: CAHOON, Rebecca, E. [US/US]; 2331 West 18th Street, Wilmington, DE 19806 (US). KINNEY, Anthony, John [GB/US]; 609 Lore Avenue, Wilmington, DE 19809 (US). RAFALSKI, J., Antoni [US/US]; 2028 Longcome Drive, Wilmington, DE 19810 (US). RENDINA, Alan, R. [US/US]; 506 Riblett Lane, Wilmington, DE 19808 (US).

(74) Agent: MAJARIAN, William, R.; E.I. Du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US).

(81) Designated States: AE, AL, AU, BA, BB, BG, BR, CA, CN, CU, CZ, EE, GD, GE, HR, HU, ID, IL, IN, IS, JP, KP, KR, LC, LK, LR, LT, LV, MG, MK, MN, MX, NO, NZ, PL, RO, SG, SI, SK, SL, TR, TT, UA, US, UZ, VN, YU, ZA, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published**
*With international search report.*
*Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.*

(54) Title: CHORISMATE BIOSYNTHESIS ENZYMES

(57) Abstract

This invention relates to an isolated nucleic acid fragment encoding a 3–dehydroquinate synthase. The invention also relates to the construction of a chimeric gene encoding all or a portion of the 3–dehydroquinate synthase, in sense or antisense orientation, wherein expression of the chimeric gene results in production of altered levels of the 3–dehydroquinate synthase in a transformed host cell.

# TITLE

## CHORISMATE BIOSYNTHESIS ENZYMES

This application claims the benefit of U.S. Provisional Application No. 60/093,611, filed July 21, 1998.

## FIELD OF THE INVENTION

This invention is in the field of plant molecular biology. More specifically, this invention pertains to nucleic acid fragments encoding enzymes involved in chorismate biosynthesis in plants and seeds.

## BACKGROUND OF THE INVENTION

Chorismate biosynthesis involves the last few steps in the common pathway for the production of the aromatic amino acids phenylalanine, tyrosine and tryptophan. The first step in chorismate biosyntheis is performed by 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase. Two differentially expressed forms of this enzyme are present in plant tissues and are important regulators of the flux to aromatic amino acid biosynthesis. The next enzyme involved in this pathway is 3-dehydroquinate synthase which has been previously described in bacteria, fungi and yeast, (Millar, G. and Coggins, J.R. (1986) *FEBS* Lett. *200*:11-17). This enzyme is encoded by the AroB gene in *E. coli* K12, its activity is dependent on NAD+ and divalent cations, and it is inhibited by the broad spectrum herbicide glyphosate. No plant 3-dehydroquinate synthase gene is known to date.

Manipulating either the amount or activity of this enzyme would afford a means to change the ratio of aromatic to non-aromatic amino acids in plants, including corn, rice, sorghum, soybean and wheat. This enzyme should also be useful for high throughput screening of compounds suitable for use as herbicides.

## SUMMARY OF THE INVENTION

The instant invention relates to isolated nucleic acid fragments encoding 3-dehydroquinate synthase. Specifically, this invention concerns an isolated nucleic acid fragment encoding a 3-dehydroquinate synthase and an isolated nucleic acid fragment that is substantially similar to an isolated nucleic acid fragment encoding a 3-dehydroquinate synthase. In addition, this invention relates to a nucleic acid fragment that is complementary to the nucleic acid fragment encoding 3-dehydroquinate synthase.

An additional embodiment of the instant invention pertains to a polypeptide encoding all or a substantial portion of a 3-dehydroquinate synthase.

In another embodiment, the instant invention relates to a chimeric gene encoding a 3-dehydroquinate synthase, or to a chimeric gene that comprises a nucleic acid fragment that is complementary to a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences, wherein expression of the chimeric gene results in production of levels of the encoded protein in a transformed host cell that is altered (i.e., increased or decreased) from the level produced in an untransformed host cell.

In a further embodiment, the instant invention concerns a transformed host cell comprising in its genome a chimeric gene encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences. Expression of the chimeric gene results in production of altered levels of the encoded protein in the transformed host cell. The transformed host cell can be of eukaryotic or prokaryotic origin, and include cells derived from higher plants and microorganisms. The invention also includes transformed plants that arise from transformed host cells of higher plants, and seeds derived from such transformed plants.

An additional embodiment of the instant invention concerns a method of altering the level of expression of a 3-dehydroquinate synthase in a transformed host cell comprising: a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase; and b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of altered levels of 3-dehydroquinate synthase in the transformed host cell.

An addition embodiment of the instant invention concerns a method for obtaining a nucleic acid fragment encoding all or a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase.

A further embodiment of the instant invention is a method for evaluating at least one compound for its ability to inhibit the activity of a 3-dehydroquinate synthase, the method comprising the steps of: (a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences; (b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of 3-dehydroquinate synthase in the transformed host cell; (c) optionally purifying the 3-dehydroquinate synthase expressed by the transformed host cell; (d) treating the 3-dehydroquinate synthase with a compound to be tested; and (e) comparing the activity of the 3-dehydroquinate synthase that has been treated with a test compound to the activity of an untreated 3-dehydroquinate synthase, thereby selecting compounds with potential for inhibitory activity.

## BRIEF DESCRIPTION OF THE DRAWING AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description and the accompanying drawing and Sequence Listing which form a part of this application.

Figure 1 depicts the amino acid sequence alignment between the 3-dehydroquinate synthase from *Escherichia coli* (NCBI General Identifier No. 114181; SEQ ID NO:9), the amino acid sequence encoded by the instant corn contig assembled from cDNA clones cr1n.pk0113.g3cbn2.pk0047.e10 and cca.pk0019.e12 (SEQ ID NO:2), the amino acid

sequence encoded by the instant rice cDNA clone rls72.pk0035.a10 (SEQ ID NO:4), the amino acid sequence encoded by the instant soybean cDNA clone se3.pk0029.f9 (SEQ ID NO:6) and the amino acid sequence encoded by the instant wheat cDNA clone wr1.pk0011.d5 (SEQ ID NO:8). The top row indicates with asterisks (*) the amino acids conserved among all sequences and with plus signs (+) the amino acids conserved only among the plant sequences. Dashes are used by the program to maximize the alignment of the sequences.

Table 1 lists the polypeptides that are described herein, the designation of the cDNA clones that comprise the nucleic acid fragments encoding polypeptides representing all or a substantial portion of these polypeptides, and the corresponding identifier (SEQ ID NO:) as used in the attached Sequence Listing. The sequence descriptions and Sequence Listing attached hereto comply with the rules governing nucleotide and/or amino acid sequence disclosures in patent applications as set forth in 37 C.F.R. §1.821-1.825.

TABLE 1

3-Dehydroquinate Synthase

| Plant | Clone Designation | SEQ ID NO: (Nucleotide) | SEQ ID NO: (Amino Acid) |
|---|---|---|---|
| Corn | Contig of: cbn2.pk0047.e10 cca.pk0019.e12 | 1 | 2 |
| Rice | rls72.pk0035.a10 | 3 | 4 |
| Soybean | se3.pk0029.f9 | 5 | 6 |
| Wheat | wr1.pk0011.d5 | 7 | 8 |

The Sequence Listing contains the one letter code for nucleotide sequence characters and the three letter codes for amino acids as defined in conformity with the IUPAC-IUBMB standards described in *Nucleic Acids Research 13*:3021-3030 (1985) and in the *Biochemical Journal 219 (No. 2)*:345-373 (1984) which are herein incorporated by reference. The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

## DETAILED DESCRIPTION OF THE INVENTION

In the context of this disclosure, a number of terms shall be utilized. As used herein, a "nucleic acid fragment" is a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. A nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

As used herein, "contig" refers to a nucleotide sequence that is assembled from two or more constituent nucleotide sequences that share common or overlapping regions of

sequence homology. For example, the nucleotide sequences of two or more nucleic acid fragments can be compared and aligned in order to identify common or overlapping sequences. Where common or overlapping sequences exist between two or more nucleic acid fragments, the sequences (and thus their corresponding nucleic acid fragments) can be assembled into a single contiguous nucleotide sequence.

As used herein, "substantially similar" refers to nucleic acid fragments wherein changes in one or more nucleotide bases results in substitution of one or more amino acids, but do not affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases does not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene silencing through for example antisense or co-suppression technology. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment which result in the production of a chemically equivalent amino acid at a given site, but do not effect the functional properties of the encoded polypeptide, are well known in the art. Thus, a codon for the amino acid alanine, a hydrophobic amino acid, may be substituted by a codon encoding another less hydrophobic residue, such as glycine, or a more hydrophobic residue, such as valine, leucine, or isoleucine. Similarly, changes which result in substitution of one negatively charged residue for another, such as aspartic acid for glutamic acid, or one positively charged residue for another, such as lysine for arginine, can also be expected to produce a functionally equivalent product. Nucleotide changes which result in alteration of the N-terminal and C-terminal portions of the polypeptide molecule would also not be expected to alter the activity of the polypeptide. Each of the proposed modifications is well within the routine skill in the art, as is determination of retention of biological activity of the encoded products.

Moreover, substantially similar nucleic acid fragments may also be characterized by their ability to hybridize. Estimates of such homology are provided by either DNA-DNA or DNA-RNA hybridization under conditions of stringency as is well understood by those skilled in the art (Hames and Higgins, Eds. (1985) *Nucleic Acid Hybridisation*, IRL Press,

Oxford, U.K.). Stringency conditions can be adjusted to screen for moderately similar fragments, such as homologous sequences from distantly related organisms, to highly similar fragments, such as genes that duplicate functional enzymes from closely related organisms. Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45°C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50°C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60°C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65°C.

Substantially similar nucleic acid fragments of the instant invention may also be characterized by the percent identity of the amino acid sequences that they encode to the amino acid sequences disclosed herein, as determined by algorithms commonly employed by those skilled in this art. Preferred are those nucleic acid fragments whose nucleotide sequences encode amino acid sequences that are 80% identical to the amino acid sequences reported herein. More preferred nucleic acid fragments encode amino acid sequences that are 90% identical to the amino acid sequences reported herein. Most preferred are nucleic acid fragments that encode amino acid sequences that are 95% identical to the amino acid sequences reported herein. Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS.* 5:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5.

A "substantial portion" of an amino acid or nucleotide sequence comprises an amino acid or a nucleotide sequence that is sufficient to afford putative identification of the protein or gene that the amino acid or nucleotide sequence comprises. Amino acid and nucleotide sequences can be evaluated either manually by one skilled in the art, or by using computer-based sequence comparison and identification tools that employ algorithms such as BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/). In general, a sequence of ten or more contiguous amino acids or thirty or more contiguous nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene-specific oligonucleotide probes comprising 30 or more contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization

of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12 or more nucleotides may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises a nucleotide sequence that will afford specific identification and/or isolation of a nucleic acid fragment comprising the sequence. The instant specification teaches amino acid and nucleotide sequences encoding polypeptides that comprise one or more particular plant proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

"Codon degeneracy" refers to divergence in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. Accordingly, the instant invention relates to any nucleic acid fragment comprising a nucleotide sequence that encodes all or a substantial portion of the amino acid sequences set forth herein. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a nucleic acid fragment for improved expression in a host cell, it is desirable to design the nucleic acid fragment such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

"Synthetic nucleic acid fragments" can be assembled from oligonucleotide building blocks that are chemically synthesized using procedures known to those skilled in the art. These building blocks are ligated and annealed to form larger nucleic acid fragments which may then be enzymatically assembled to construct the entire desired nucleic acid fragment. "Chemically synthesized", as related to nucleic acid fragment, means that the component nucleotides were assembled *in vitro*. Manual chemical synthesis of nucleic acid fragments may be accomplished using well established procedures, or automated chemical synthesis can be performed using one of a number of commercially available machines. Accordingly, the nucleic acid fragments can be tailored for optimal gene expression based on optimization of nucleotide sequence to reflect the codon bias of the host cell. The skilled artisan appreciates the likelihood of successful gene expression if codon usage is biased towards those codons favored by the host. Determination of preferred codons can be based on a survey of genes derived from the host cell where sequence information is available.

"Gene" refers to a nucleic acid fragment that expresses a specific protein, including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature.

Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign" gene refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

"Coding sequence" refers to a nucleotide sequence that codes for a specific amino acid sequence. "Regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, and polyadenylation recognition sequences.

"Promoter" refers to a nucleotide sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. The promoter sequence consists of proximal and more distal upstream elements, the latter elements often referred to as enhancers. Accordingly, an "enhancer" is a nucleotide sequence which can stimulate promoter activity and may be an innate element of the promoter or a heterologous element inserted to enhance the level or tissue-specificity of a promoter. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleotide segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental conditions. Promoters which cause a nucleic acid fragment to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". New promoters of various types useful in plant cells are constantly being discovered; numerous examples may be found in the compilation by Okamuro and Goldberg (1989) *Biochemistry of Plants 15*:1-82. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, nucleic acid fragments of different lengths may have identical promoter activity.

The "translation leader sequence" refers to a nucleotide sequence located between the promoter sequence of a gene and the coding sequence. The translation leader sequence is present in the fully processed mRNA upstream of the translation start sequence. The translation leader sequence may affect processing of the primary transcript to mRNA, mRNA stability or translation efficiency. Examples of translation leader sequences have been described (Turner and Foster (1995) *Molecular Biotechnology 3*:225).

The "3' non-coding sequences" refer to nucleotide sequences located downstream of a coding sequence and include polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The use of different 3' non-coding sequences is exemplified by Ingelbrecht et al. (1989) *Plant Cell 1*:671-680.

"RNA transcript" refers to the product resulting from RNA polymerase-catalyzed transcription of a DNA sequence. When the RNA transcript is a perfect complementary copy of the DNA sequence, it is referred to as the primary transcript or it may be a RNA sequence derived from posttranscriptional processing of the primary transcript and is referred to as the mature RNA. "Messenger RNA (mRNA)" refers to the RNA that is without introns and that can be translated into polypeptide by the cell. "cDNA" refers to a double-stranded DNA that is complementary to and derived from mRNA. "Sense" RNA refers to an RNA transcript that includes the mRNA and so can be translated into a polypeptide by the cell. "Antisense RNA" refers to an RNA transcript that is complementary to all or part of a target primary transcript or mRNA and that blocks the expression of a target gene (see U.S. Patent No. 5,107,065, incorporated herein by reference). The complementarity of an antisense RNA may be with any part of the specific nucleotide sequence, i.e., at the 5' non-coding sequence, 3' non-coding sequence, introns, or the coding sequence. "Functional RNA" refers to sense RNA, antisense RNA, ribozyme RNA, or other RNA that may not be translated but yet has an effect on cellular processes.

The term "operably linked" refers to the association of two or more nucleic acid fragments on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide. "Antisense inhibition" refers to the production of antisense RNA transcripts capable of suppressing the expression of the target protein. "Overexpression" refers to the production of a gene product in transgenic organisms that exceeds levels of production in normal or non-transformed organisms. "Co-suppression" refers to the production of sense RNA transcripts capable of suppressing the expression of identical or substantially similar foreign or endogenous genes (U.S. Patent No. 5,231,020, incorporated herein by reference).

"Altered levels" refers to the production of gene product(s) in transgenic organisms in amounts or proportions that differ from that of normal or non-transformed organisms.

"Mature" protein refers to a post-translationally processed polypeptide; i.e., one from which any pre- or propeptides present in the primary translation product have been removed. "Precursor" protein refers to the primary product of translation of mRNA; i.e., with pre- and propeptides still present. Pre- and propeptides may be but are not limited to intracellular localization signals.

"Transformation" refers to the transfer of a nucleic acid fragment into the genome of a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" organisms. Examples of methods of plant transformation include *Agrobacterium*-mediated transformation (De Blaere et al. (1987) *Meth. Enzymol. 143*:277) and particle-accelerated or "gene gun" transformation technology (Klein et al. (1987) *Nature (London) 327*:70-73; U.S. Patent No. 4,945,050, incorporated herein by reference).

Standard recombinant DNA and molecular cloning techniques used herein are well known in the art and are described more fully in Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, 1989 (hereinafter "Maniatis").

Nucleic acid fragments encoding at least a portion of several 3-dehydroquinate synthases have been isolated and identified by comparison of random plant cDNA sequences to public databases containing nucleotide and protein sequences using the BLAST algorithms well known to those skilled in the art. The nucleic acid fragments of the instant invention may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to, methods of nucleic acid hybridization, and methods of DNA and RNA amplification as exemplified by various uses of nucleic acid amplification technologies (e.g., polymerase chain reaction, ligase chain reaction).

For example, genes encoding other 3-dehydroquinate synthases, either as cDNAs or genomic DNAs, could be isolated directly by using all or a portion of the instant nucleic acid fragments as DNA hybridization probes to screen libraries from any desired plant employing methodology well known to those skilled in the art. Specific oligonucleotide probes based upon the instant nucleic acid sequences can be designed and synthesized by methods known in the art (Maniatis). Moreover, the entire sequences can be used directly to synthesize DNA probes by methods known to the skilled artisan such as random primer DNA labeling, nick translation, or end-labeling techniques, or RNA probes using available *in vitro* transcription systems. In addition, specific primers can be designed and used to amplify a part or all of the instant sequences. The resulting amplification products can be labeled directly during amplification reactions or labeled after amplification reactions, and used as

probes to isolate full length cDNA or genomic fragments under conditions of appropriate stringency.

In addition, two short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. The polymerase chain reaction may also be performed on a library of cloned nucleic acid fragments wherein the sequence of one primer is derived from the instant nucleic acid fragments, and the sequence of the other primer takes advantage of the presence of the polyadenylic acid tracts to the 3' end of the mRNA precursor encoding plant genes. Alternatively, the second primer sequence may be based upon sequences derived from the cloning vector. For example, the skilled artisan can follow the RACE protocol (Frohman et al. (1988) *Proc. Natl. Acad. Sci. USA 85*:8998) to generate cDNAs by using PCR to amplify copies of the region between a single point in the transcript and the 3' or 5' end. Primers oriented in the 3' and 5' directions can be designed from the instant sequences. Using commercially available 3' RACE or 5' RACE systems (BRL), specific 3' or 5' cDNA fragments can be isolated (Ohara et al. (1989) *Proc. Natl. Acad. Sci. USA 86*:5673; Loh et al. (1989) *Science 243*:217). Products generated by the 3' and 5' RACE procedures can be combined to generate full-length cDNAs (Frohman and Martin (1989) *Techniques 1*:165).

Availability of the instant nucleotide and deduced amino acid sequences facilitates immunological screening of cDNA expression libraries. Synthetic peptides representing portions of the instant amino acid sequences may be synthesized. These peptides can be used to immunize animals to produce polyclonal or monoclonal antibodies with specificity for peptides or proteins comprising the amino acid sequences. These antibodies can be then be used to screen cDNA expression libraries to isolate full-length cDNA clones of interest (Lerner (1984) *Adv. Immunol. 36*:1; Maniatis).

The nucleic acid fragments of the instant invention may be used to create transgenic plants in which the disclosed polypeptides are present at higher or lower levels than normal or in cell types or developmental stages in which they are not normally found. This would have the effect of altering the ratio of aromatic to non-aromatic amino acids in those cells. This enzyme may also be used for the high throughput screening of compounds suitable for use as herbicides.

Overexpression of the proteins of the instant invention may be accomplished by first constructing a chimeric gene in which the coding region is operably linked to a promoter capable of directing expression of a gene in the desired tissues at the desired stage of development. For reasons of convenience, the chimeric gene may comprise promoter sequences and translation leader sequences derived from the same genes. 3' Non-coding sequences encoding transcription termination signals may also be provided. The instant chimeric gene may also comprise one or more introns in order to facilitate gene expression.

Plasmid vectors comprising the instant chimeric gene can then constructed. The choice of plasmid vector is dependent upon the method that will be used to transform host plants. The skilled artisan is well aware of the genetic elements that must be present on the plasmid vector in order to successfully transform, select and propagate host cells containing the chimeric gene. The skilled artisan will also recognize that different independent transformation events will result in different levels and patterns of expression (Jones et al. (1985) *EMBO J. 4*:2411-2418; De Almeida et al. (1989) *Mol. Gen. Genetics 218*:78-86), and thus that multiple events must be screened in order to obtain lines displaying the desired expression level and pattern. Such screening may be accomplished by Southern analysis of DNA, Northern analysis of mRNA expression, Western analysis of protein expression, or phenotypic analysis.

It may also be desirable to reduce or eliminate expression of genes encoding the instant polypeptide in plants for some applications. In order to accomplish this, a chimeric gene designed for co-suppression of the instant polypeptide can be constructed by linking a gene or gene fragment encoding that polypeptide to plant promoter sequences. Alternatively, a chimeric gene designed to express antisense RNA for all or part of the instant nucleic acid fragment can be constructed by linking the gene or gene fragment in reverse orientation to plant promoter sequences. Either the co-suppression or antisense chimeric genes could be introduced into plants via transformation wherein expression of the corresponding endogenous genes are reduced or eliminated.

Molecular genetic solutions to the generation of plants with altered gene expression have a decided advantage over more traditional plant breeding approaches. Changes in plant phenotypes can be produced by specifically inhibiting expression of one or more genes by antisense inhibition or cosuppression (U. S. Patent Nos. 5,190,931, 5,107,065 and 5,283,323). An antisense or cosuppression construct would act as a dominant negative regulator of gene activity. While conventional mutations can yield negative regulation of gene activity these effects are most likely recessive. The dominant negative regulation available with a transgenic approach may be advantageous from a breeding perspective. In addition, the ability to restrict the expression of specific phenotype to the reproductive tissues of the plant by the use of tissue specific promoters may confer agronomic advantages relative to conventional mutations which may have an effect in all tissues in which a mutant gene is ordinarily expressed.

The person skilled in the art will know that special considerations are associated with the use of antisense or cosuppresion technologies in order to reduce expression of particular genes. For example, the proper level of expression of sense or antisense genes may require the use of different chimeric genes utilizing different regulatory elements known to the skilled artisan. Once transgenic plants are obtained by one of the methods described above, it will be necessary to screen individual transgenics for those that most effectively display

the desired phenotype. Accordingly, the skilled artisan will develop methods for screening large numbers of transformants. The nature of these screens will generally be chosen on practical grounds, and is not an inherent part of the invention. For example, one can screen by looking for changes in gene expression by using antibodies specific for the protein encoded by the gene being suppressed, or one could establish assays that specifically measure enzyme activity. A preferred method will be one which allows large numbers of samples to be processed rapidly, since it will be expected that a large number of transformants will be negative for the desired phenotype.

The instant polypeptide (or portions thereof) may be produced in heterologous host cells, particularly in the cells of microbial hosts, and can be used to prepare antibodies to the these proteins by methods well known to those skilled in the art. The antibodies are useful for detecting the polypeptide of the instant invention *in situ* in cells or *in vitro* in cell extracts. Preferred heterologous host cells for production of the instant polypeptide are microbial hosts. Microbial expression systems and expression vectors containing regulatory sequences that direct high level expression of foreign proteins are well known to those skilled in the art. Any of these could be used to construct a chimeric gene for production of the instant polypeptide. This chimeric gene could then be introduced into appropriate microorganisms via transformation to provide high level expression of the encoded 3-dehydroquinate synthase. An example of a vector for high level expression of the instant polypeptide in a bacterial host is provided (Example 6).

Additionally, the instant polypeptide can be used as a target to facilitate design and/or identification of inhibitors of those enzymes that may be useful as herbicides. This is desirable because the polypeptide described herein catalyzes a step in aromatic amino acid biosynthesis. Accordingly, inhibition of the activity of one or more of the enzymes described herein could lead to inhibition plant growth. Thus, the instant polypeptide could be appropriate for new herbicide discovery and design.

All or a substantial portion of the nucleic acid fragments of the instant invention may also be used as probes for genetically and physically mapping the genes that they are a part of, and as markers for traits linked to those genes. Such information may be useful in plant breeding in order to develop lines with desired phenotypes. For example, the instant nucleic acid fragments may be used as restriction fragment length polymorphism (RFLP) markers. Southern blots (Maniatis) of restriction-digested plant genomic DNA may be probed with the nucleic acid fragments of the instant invention. The resulting banding patterns may then be subjected to genetic analyses using computer programs such as MapMaker (Lander et al. (1987) *Genomics 1*:174-181) in order to construct a genetic map. In addition, the nucleic acid fragments of the instant invention may be used to probe Southern blots containing restriction endonuclease-treated genomic DNAs of a set of individuals representing parent and progeny of a defined genetic cross. Segregation of the DNA polymorphisms is noted

and used to calculate the position of the instant nucleic acid sequence in the genetic map previously obtained using this population (Botstein et al. (1980) *Am. J. Hum. Genet. 32*:314-331).

The production and use of plant gene-derived probes for use in genetic mapping is described in Bernatzky and Tanksley (1986) *Plant Mol. Biol. Reporter 4(1)*:37-41. Numerous publications describe genetic mapping of specific cDNA clones using the methodology outlined above or variations thereof. For example, F2 intercross populations, backcross populations, randomly mated populations, near isogenic lines, and other sets of individuals may be used for mapping. Such methodologies are well known to those skilled in the art.

Nucleic acid probes derived from the instant nucleic acid sequences may also be used for physical mapping (i.e., placement of sequences on physical maps; *see* Hoheisel et al. In: *Nonmammalian Genomic Analysis: A Practical Guide*, Academic press 1996, pp. 319-346, and references cited therein).

In another embodiment, nucleic acid probes derived from the instant nucleic acid sequences may be used in direct fluorescence *in situ* hybridization (FISH) mapping (Trask (1991) *Trends Genet. 7*:149-154). Although current methods of FISH mapping favor use of large clones (several to several hundred KB; see Laan et al. (1995) *Genome Research 5*:13-20), improvements in sensitivity may allow performance of FISH mapping using shorter probes.

A variety of nucleic acid amplification-based methods of genetic and physical mapping may be carried out using the instant nucleic acid sequences. Examples include allele-specific amplification (Kazazian (1989) *J. Lab. Clin. Med. 114(2)*:95-96), polymorphism of PCR-amplified fragments (CAPS; Sheffield et al. (1993) *Genomics 16*:325-332), allele-specific ligation (Landegren et al. (1988) *Science 241*:1077-1080), nucleotide extension reactions (Sokolov (1990) *Nucleic Acid Res. 18*:3671), Radiation Hybrid Mapping (Walter et al. (1997) *Nature Genetics 7*:22-28) and Happy Mapping (Dear and Cook (1989) *Nucleic Acid Res. 17*:6795-6807). For these methods, the sequence of a nucleic acid fragment is used to design and produce primer pairs for use in the amplification reaction or in primer extension reactions. The design of such primers is well known to those skilled in the art. In methods employing PCR-based genetic mapping, it may be necessary to identify DNA sequence differences between the parents of the mapping cross in the region corresponding to the instant nucleic acid sequence. This, however, is generally not necessary for mapping methods.

Loss of function mutant phenotypes may be identified for the instant cDNA clones either by targeted gene disruption protocols or by identifying specific mutants for these genes contained in a maize population carrying mutations in all possible genes (Ballinger and Benzer (1989) *Proc. Natl. Acad. Sci USA 86*:9402; Koes et al. (1995) *Proc. Natl. Acad.*

*Sci USA 92*:8149; Bensen et al. (1995) *Plant Cell 7*:75). The latter approach may be accomplished in two ways. First, short segments of the instant nucleic acid fragments may be used in polymerase chain reaction protocols in conjunction with a mutation tag sequence primer on DNAs prepared from a population of plants in which Mutator transposons or some other mutation-causing DNA element has been introduced (see Bensen, *supra*). The amplification of a specific DNA fragment with these primers indicates the insertion of the mutation tag element in or near the plant gene encoding the instant polypeptide. Alternatively, the instant nucleic acid fragment may be used as a hybridization probe against PCR amplification products generated from the mutation population using the mutation tag sequence primer in conjunction with an arbitrary genomic site primer, such as that for a restriction enzyme site-anchored synthetic adaptor. With either method, a plant containing a mutation in the endogenous gene encoding the instant polypeptide can be identified and obtained. This mutant plant can then be used to determine or confirm the natural function of the instant polypeptide disclosed herein.

## EXAMPLES

The present invention is further defined in the following Examples, in which all parts and percentages are by weight and degrees are Celsius, unless otherwise stated. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various usages and conditions.

### EXAMPLE 1

#### Composition of cDNA Libraries; Isolation and Sequencing of cDNA Clones

cDNA libraries representing mRNAs from various corn, rice, soybean and wheat tissues were prepared. The characteristics of the libraries are described below.

TABLE 2

cDNA Libraries from Corn, Rice, Soybean and Wheat

| Library | Tissue | Clone |
|---|---|---|
| cbn2 | Corn Developing Kernel Two Days After Pollination | cbn2.pk0047.e10 |
| cca | Corn Callus Type II Tissue, Undifferentiated, Highly Transformable | cca.pk0019.e12 |
| rls72 | Rice Leaf 15 Days After Germination, 72 Hours After Infection of Strain *Magaporthe grisea* 4360-R-67 (AVR2-YAMO); Susceptible | rls72.pk0035.a10 |
| se3 | Soybean Embryo, 17 Days After Flowering | se3.pk0029.f9 |
| wr1 | Wheat Root From 7 Day Old Seedling | wr1.pk0011.d5 |

cDNA libraries may be prepared by any one of many methods available. For example, the cDNAs may be introduced into plasmid vectors by first preparing the cDNA libraries in Uni-ZAP* XR vectors according to the manufacturer's protocol (Stratagene Cloning Systems, La Jolla, CA). The Uni-ZAP* XR libraries are converted into plasmid libraries according to the protocol provided by Stratagene. Upon conversion, cDNA inserts will be contained in the plasmid vector pBluescript. In addition, the cDNAs may be introduced directly into precut Bluescript II SK(+) vectors (Stratagene) using T4 DNA ligase (New England Biolabs), followed by transfection into DH10B cells according to the manufacturer's protocol (GIBCO BRL Products). Once the cDNA inserts are in plasmid vectors, plasmid DNAs are prepared from randomly picked bacterial colonies containing recombinant pBluescript plasmids, or the insert cDNA sequences are amplified via polymerase chain reaction using primers specific for vector sequences flanking the inserted cDNA sequences. Amplified insert DNAs or plasmid DNAs are sequenced in dye-primer sequencing reactions to generate partial cDNA sequences (expressed sequence tags or "ESTs"; see Adams et al., (1991) *Science 252*:1651). The resulting ESTs are analyzed using a Perkin Elmer Model 377 fluorescent sequencer.

## EXAMPLE 2

### Identification of cDNA Clones

cDNA clones encoding 3-dehydroquinate synthase were identified by conducting BLAST (Basic Local Alignment Search Tool; Altschul et al. (1993) *J. Mol. Biol. 215*:403-410; see also www.ncbi.nlm.nih.gov/BLAST/) searches for similarity to sequences contained in the BLAST "nr" database (comprising all non-redundant GenBank CDS translations, sequences derived from the 3-dimensional structure Brookhaven Protein Data Bank, the last major release of the SWISS-PROT protein sequence database, EMBL, and DDBJ databases). The cDNA sequences obtained in Example 1 were analyzed for similarity to all publicly available DNA sequences contained in the "nr" database using the BLASTN algorithm provided by the National Center for Biotechnology Information (NCBI). The DNA sequences were translated in all reading frames and compared for similarity to all publicly available protein sequences contained in the "nr" database using the BLASTX algorithm (Gish and States (1993) *Nature Genetics 3*:266-272) provided by the NCBI. For convenience, the P-value (probability) of observing a match of a cDNA sequence to a sequence contained in the searched databases merely by chance as calculated by BLAST are reported herein as "pLog" values, which represent the negative of the logarithm of the reported P-value. Accordingly, the greater the pLog value, the greater the likelihood that the cDNA sequence and the BLAST "hit" represent homologous proteins.

EXAMPLE 3

Characterization of cDNA Clones Encoding 3-Dehydroquinate Synthase

The BLASTX search using the EST sequences from clones listed in Table 3 revealed similarity of the polypeptides encoded by the cDNAs to 3-dehydroquinate synthase from *Escherichia coli* (NCBI General Identifier No. 114181). Shown in Table 3 are the BLAST results for the entire cDNA inserts comprising the indicated cDNA clones ("FIS"), or contigs assembled from an FIS and an EST ("Contig"):

TABLE 3

BLAST Results for Sequences Encoding Polypeptides Homologous to 3-Dehydroquinate Synthase

| Clone | Status | BLAST pLog Score 114181 |
|---|---|---|
| Contig of:<br>cbn2.pk0047.e10<br>cca.pk0019.e12 | Contig | 102.00 |
| rls72.pk0035.a10 | FIS | 102.00 |
| se3.pk0029.f9 | FIS | 101.00 |
| wr1.pk0011.d5 | FIS | 79.00 |

BLASTN analysis of the NCBI EST database indicates that nucleotides 956 to 1300 of the contig assembled from clones cbn2.pk0047.e10 and cca.pk0019.e12 are 99% identical to nucleotides 345 to 1 of a 345 nucleotide corn EST (NCBI General Identifier No. 4688530). The EST was published in NCBI on April 26, 1999. Clone cca.pk0019.e12 encodes the entire 3-dehydroquinate synthase gene but is incompletely processed. Using clone cbn2.pk0047.e10, the sequence corresponding to the intron has been removed. The instant corn 3-dehydroquinate synthase sequence was determined prior to publication of the corn EST in NCBI.

BLASTN analysis of the NCBI EST database indicates that nucleotides 1005 to 1328 from clone rls72.pk0035.a10 are 96% identical to nucleotides 20 to 344 of a 683 nucleotide rice EST (NCBI General Identifier No. 4715434) published on April 29, 1999. Clone rls72.pk0035.a10 encodes an entire 3-dehydroquinate synthase; this sequence was determined prior to publication of the rice EST sequence.

Figure 1 presents an alignment of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6 and 8 and the *Escherichia coli* sequence (SEQ ID NO:9). The data in Table 4 represents a calculation of the percent identity of the amino acid sequences set forth in SEQ ID NOs:2, 4, 6 and 8 and the *Escherichia coli* sequence (SEQ ID NO:9).

TABLE 4

Percent Identity of Amino Acid Sequences Deduced From the Nucleotide Sequences of cDNA Clones Encoding Polypeptides Homologous to 3-Dehydroquinate Synthase

| SEQ ID NO. | Percent Identity to 114181 |
|---|---|
| 2 | 52.5 |
| 4 | 52.5 |
| 6 | 52.8 |
| 8 | 55.0 |

Sequence alignments and percent identity calculations were performed using the Megalign program of the LASARGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences was performed using the Clustal method of alignment (Higgins and Sharp (1989) *CABIOS. 5*:151-153) with the default parameters (GAP PENALTY=10, GAP LENGTH PENALTY=10). Default parameters for pairwise alignments using the Clustal method were KTUPLE 1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. Sequence alignments and BLAST scores and probabilities indicate that the nucleic acid fragments comprising the instant cDNA clones encode a substantial portion of a 3-dehydroquinate synthase. These sequences represent the first plant sequences encoding 3-dehydroquinate synthase.

EXAMPLE 4

Expression of Chimeric Genes in Monocot Cells

A chimeric gene comprising a cDNA encoding the instant polypeptide in sense orientation with respect to the maize 27 kD zein promoter that is located 5' to the cDNA fragment, and the 10 kD zein 3' end that is located 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites (NcoI or SmaI) can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector pML103 as described below. Amplification is then performed in a standard PCR. The amplified DNA is then digested with restriction enzymes NcoI and SmaI and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with a 4.9 kb NcoI-SmaI fragment of the plasmid pML103. Plasmid pML103 has been deposited under the terms of the Budapest Treaty at ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209), and bears accession number ATCC 97366. The DNA segment from pML103 contains a 1.05 kb SalI-NcoI promoter fragment of the maize 27 kD zein gene and a 0.96 kb SmaI-SalI fragment from the 3' end of the maize 10 kD zein gene in the vector pGem9Zf(+) (Promega). Vector and insert DNA can be ligated at 15°C overnight,

essentially as described (Maniatis). The ligated DNA may then be used to transform *E. coli* XL1-Blue (Epicurian Coli XL-1 Blue™; Stratagene). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA Sequencing Kit; U.S. Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding, in the 5' to 3' direction, the maize 27 kD zein promoter, a cDNA fragment encoding the instant polypeptide, and the 10 kD zein 3' region.

The chimeric gene described above can then be introduced into corn cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines H99 and LH132. The embryos are isolated 10 to 11 days after pollination when they are 1.0 to 1.5 mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin. Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The plasmid, p35S/Ac (obtained from Dr. Peter Eckes, Hoechst Ag, Frankfurt, Germany) may be used in transformation experiments in order to provide for a selectable marker. This plasmid contains the *Pat* gene (see European Patent Publication 0 242 236) which encodes phosphinothricin acetyl transferase (PAT). The enzyme PAT confers resistance to herbicidal glutamine synthetase inhibitors such as phosphinothricin. The *pat* gene in p35S/Ac is under the control of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature* 313:810-812) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1 μm in diameter) are coated with DNA using the following technique. Ten μg of plasmid DNAs are added to 50 μL of a suspension of gold particles (60 mg per mL). Calcium chloride (50 μL of a 2.5 M solution) and spermidine free base (20 μL of a 1.0 M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 μL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 μL of ethanol. An aliquot (5 μL) of the DNA-coated gold particles can be placed in the center of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad

Instruments. Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment. the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardment the tissue can be transferred to N6 medium that contains gluphosinate (2 mg per liter) and lacks casein or proline. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing gluphosinate. After 6 weeks. areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the glufosinate-supplemented medium. These calli may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et al. (1990) *Bio/Technology* *8*:833-839).

## EXAMPLE 5

### Expression of Chimeric Genes in Dicot Cells

A seed-specific expression cassette composed of the promoter and transcription terminator from the gene encoding the β subunit of the seed storage protein phaseolin from the bean *Phaseolus vulgaris* (Doyle et al. (1986) *J. Biol. Chem.* *261*:9228-9238) can be used for expression of the instant polypeptide in transformed soybean. The phaseolin cassette includes about 500 nucleotides upstream (5') from the translation initiation codon and about 1650 nucleotides downstream (3') from the translation stop codon of phaseolin. Between the 5' and 3' regions are the unique restriction endonuclease sites Nco I (which includes the ATG translation initiation codon), Sma I, Kpn I and Xba I. The entire cassette is flanked by Hind III sites.

The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the expression vector. Amplification is then performed as described above, and the isolated fragment is inserted into a pUC18 vector carrying the seed expression cassette.

Soybean embroys may then be transformed with the expression vector comprising sequences encoding the instant polypeptide. To induce somatic embryos, cotyledons, 3-5 mm in length dissected from surface sterilized, immature seeds of the soybean cultivar A2872, can be cultured in the light or dark at 26°C on an appropriate agar medium for 6-10 weeks. Somatic embryos which produce secondary embryos are then excised and placed into a suitable liquid medium. After repeated selection for clusters of somatic embryos which multiplied as early, globular staged embryos, the suspensions are maintained as described below.

Soybean embryogenic suspension cultures can maintained in 35 mL liquid media on a rotary shaker, 150 rpm, at 26°C with florescent lights on a 16:8 hour day/night schedule. Cultures are subcultured every two weeks by inoculating approximately 35 mg of tissue into 35 mL of liquid medium.

Soybean embryogenic suspension cultures may then be transformed by the method of particle gun bombardment (Klein et al. (1987) *Nature* (London) *327*:70, U.S. Patent No. 4,945,050). A DuPont Biolistic™ PDS1000/HE instrument (helium retrofit) can be used for these transformations.

A selectable marker gene which can be used to facilitate soybean transformation is a chimeric gene composed of the 35S promoter from Cauliflower Mosaic Virus (Odell et al. (1985) *Nature 313*:810-812), the hygromycin phosphotransferase gene from plasmid pJR225 (from *E. coli*; Gritz et al.(1983) *Gene 25*:179-188) and the 3' region of the nopaline synthase gene from the T-DNA of the Ti plasmid of *Agrobacterium tumefaciens*. The seed expression cassette comprising the phaseolin 5' region, the fragment encoding the instant polypeptide and the phaseolin 3' region can be isolated as a restriction fragment. This fragment can then be inserted into a unique restriction site of the vector carrying the marker gene.

To 50 μL of a 60 mg/mL 1 μm gold particle suspension is added (in order): 5 μL DNA (1 μg/μL), 20 μl spermidine (0.1 M), and 50 μL $CaCl_2$ (2.5 M). The particle preparation is then agitated for three minutes, spun in a microfuge for 10 seconds and the supernatant removed. The DNA-coated particles are then washed once in 400 μL 70% ethanol and resuspended in 40 μL of anhydrous ethanol. The DNA/particle suspension can be sonicated three times for one second each. Five μL of the DNA-coated gold particles are then loaded on each macro carrier disk.

Approximately 300-400 mg of a two-week-old suspension culture is placed in an empty 60x15 mm petri dish and the residual liquid removed from the tissue with a pipette. For each transformation experiment, approximately 5-10 plates of tissue are normally bombarded. Membrane rupture pressure is set at 1100 psi and the chamber is evacuated to a vacuum of 28 inches mercury. The tissue is placed approximately 3.5 inches away from the retaining screen and bombarded three times. Following bombardment, the tissue can be divided in half and placed back into liquid and cultured as described above.

Five to seven days post bombardment, the liquid media may be exchanged with fresh media, and eleven to twelve days post bombardment with fresh media containing 50 mg/mL hygromycin. This selective media can be refreshed weekly. Seven to eight weeks post bombardment, green, transformed tissue may be observed growing from untransformed, necrotic embryogenic clusters. Isolated green tissue is removed and inoculated into individual flasks to generate new, clonally propagated, transformed embryogenic suspension cultures. Each new line may be treated as an independent transformation event. These suspensions can then be subcultured and maintained as clusters of immature embryos or regenerated into whole plants by maturation and germination of individual somatic embryos.

## EXAMPLE 6

### Expression of Chimeric Genes in Microbial Cells

The cDNAs encoding the instant polypeptide can be inserted into the T7 *E. coli* expression vector pBT430. This vector is a derivative of pET-3a (Rosenberg et al. (1987) *Gene 56*:125-135) which employs the bacteriophage T7 RNA polymerase/T7 promoter system. Plasmid pBT430 was constructed by first destroying the EcoR I and Hind III sites in pET-3a at their original positions. An oligonucleotide adaptor containing EcoR I and Hind III sites was inserted at the BamH I site of pET-3a. This created pET-3aM with additional unique cloning sites for insertion of genes into the expression vector. Then, the Nde I site at the position of translation initiation was converted to an Nco I site using oligonucleotide-directed mutagenesis. The DNA sequence of pET-3aM in this region, 5'-CATATGG, was converted to 5'-CCCATGG in pBT430.

Plasmid DNA containing a cDNA may be appropriately digested to release a nucleic acid fragment encoding the protein. This fragment may then be purified on a 1% NuSieve GTG™ low melting agarose gel (FMC). Buffer and agarose contain 10 μg/ml ethidium bromide for visualization of the DNA fragment. The fragment can then be purified from the agarose gel by digestion with GELase™ (Epicentre Technologies) according to the manufacturer's instructions, ethanol precipitated, dried and resuspended in 20 μL of water. Appropriate oligonucleotide adapters may be ligated to the fragment using T4 DNA ligase (New England Biolabs, Beverly, MA). The fragment containing the ligated adapters can be purified from the excess adapters using low melting agarose as described above. The vector pBT430 is digested, dephosphorylated with alkaline phosphatase (NEB) and deproteinized with phenol/chloroform as described above. The prepared vector pBT430 and fragment can then be ligated at 16°C for 15 hours followed by transformation into DH5 electrocompetent cells (GIBCO BRL). Transformants can be selected on agar plates containing LB media and 100 μg/mL ampicillin. Transformants containing the gene encoding the instant polypeptide are then screened for the correct orientation with respect to the T7 promoter by restriction enzyme analysis.

For high level expression, a plasmid clone with the cDNA insert in the correct orientation relative to the T7 promoter can be transformed into *E. coli* strain BL21(DE3) (Studier et al. (1986) *J. Mol. Biol. 189*:113-130). Cultures are grown in LB medium containing ampicillin (100 mg/L) at 25°C. At an optical density at 600 nm of approximately 1, IPTG (isopropylthio-β-galactoside, the inducer) can be added to a final concentration of 0.4 mM and incubation can be continued for 3 h at 25°. Cells are then harvested by centrifugation and re-suspended in 50 μL of 50 mM Tris-HCl at pH 8.0 containing 0.1 mM DTT and 0.2 mM phenyl methylsulfonyl fluoride. A small amount of 1 mm glass beads can be added and the mixture sonicated 3 times for about 5 seconds each time with a microprobe sonicator. The mixture is centrifuged and the protein concentration of the supernatant determined. One μg of protein from the soluble fraction of the culture can be separated by SDS-polyacrylamide gel electrophoresis. Gels can be observed for protein bands migrating at the expected molecular weight.

## EXAMPLE 7

## Evaluating Compounds for Their Ability to Inhibit the Activity of 3-Dehydroquinate Synthase

The polypeptide described herein may be produced using any number of methods known to those skilled in the art. Such methods include, but are not limited to, expression in bacteria as described in Example 6, or expression in eukaryotic cell culture, *in planta*, and using viral expression systems in suitably infected organisms or cell lines. The instant polypeptide may be expressed either as mature forms of the proteins as observed *in vivo* or as fusion proteins by covalent attachment to a variety of enzymes, proteins or affinity tags. Common fusion protein partners include glutathione S-transferase ("GST"), thioredoxin ("Trx"), maltose binding protein, and C- and/or N-terminal hexahistidine polypeptide ("$(His)_6$"). The fusion proteins may be engineered with a protease recognition site at the fusion point so that fusion partners can be separated by protease digestion to yield intact mature enzyme. Examples of such proteases include thrombin, enterokinase and factor Xa. However, any protease can be used which specifically cleaves the peptide connecting the fusion protein and the enzyme.

Purification of the instant polypeptide, if desired, may utilize any number of separation technologies familiar to those skilled in the art of protein purification. Examples of such methods include, but are not limited to, homogenization, filtration, centrifugation, heat denaturation, ammonium sulfate precipitation, desalting, pH precipitation, ion exchange chromatography, hydrophobic interaction chromatography and affinity chromatography, wherein the affinity ligand represents a substrate, substrate analog or inhibitor. When the instant polypeptide are expressed as fusion proteins, the purification protocol may include the use of an affinity resin which is specific for the fusion protein tag attached to the expressed enzyme or an affinity resin containing ligands which are specific for the enzyme.

For example, the instant polypeptide may be expressed as a fusion protein coupled to the C-terminus of thioredoxin. In addition, a $(His)_6$ peptide may be engineered into the N-terminus of the fused thioredoxin moiety to afford additional opportunities for affinity purification. Other suitable affinity resins could be synthesized by linking the appropriate ligands to any suitable resin such as Sepharose-4B. In an alternate embodiment, a thioredoxin fusion protein may be eluted using dithiothreitol; however, elution may be accomplished using other reagents which interact to displace the thioredoxin from the resin. These reagents include β–mercaptoethanol or other reduced thiol. The eluted fusion protein may be subjected to further purification by traditional means as stated above, if desired. Proteolytic cleavage of the thioredoxin fusion protein and the enzyme may be accomplished after the fusion protein is purified or while the protein is still bound to the ThioBond™ affinity resin or other resin.

Crude, partially purified or purified enzyme, either alone or as a fusion protein, may be utilized in assays for the evaluation of compounds for their ability to inhibit enzymatic activation of the instant polypeptide disclosed herein. Assays may be conducted under well known experimental conditions which permit optimal enzymatic activity. Assays for 3-dehydroquinate synthase are presented by Mehdi et al. (1987) *Methods Enzymol. 142*:306-314; Bender et al. (1989) *Biochemistry 28*:7555-7560 and Gollub et al. (1971) *Methods Enzymol. 17A*:349.

## CLAIMS

What is claimed is:

1. An isolated nucleic acid fragment encoding a 3-dehydroquinate synthase comprising a member selected from the group consisting of:

(a) an isolated nucleic acid fragment encoding an amino acid sequence that is at least 80% identical to the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 and SEQ ID NO:8;

(b) an isolated nucleic acid fragment that is complementary to (a).

2. The isolated nucleic acid fragment of Claim 1 wherein nucleic acid fragment is a functional RNA.

3. The isolated nucleic acid fragment of Claim 1 wherein the nucleotide sequence of the fragment comprises the sequence set forth in a member selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 and SEQ ID NO:7.

4. A chimeric gene comprising the nucleic acid fragment of Claim 1 operably linked to suitable regulatory sequences.

5. A transformed host cell comprising the chimeric gene of Claim 4.

6. A 3-dehydroquinate synthase polypeptide comprising all or a substantial portion of the amino acid sequence set forth in a member selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 and SEQ ID NO:8.

7. A method of altering the level of expression of a 3-dehydroquinate synthase in a host cell comprising:

(a) transforming a host cell with the chimeric gene of Claim 4 and

(b) growing the transformed host cell produced in step (a) under conditions that are suitable for expression of the chimeric gene

wherein expression of the chimeric gene results in production of altered levels of a 3-dehydroquinate synthase in the transformed host cell.

8. A method of obtaining a nucleic acid fragment encoding all or a substantial portion of the amino acid sequence encoding a 3-dehydroquinate synthase comprising:

(a) probing a cDNA or genomic library with the nucleic acid fragment of Claim 1;

(b) identifying a DNA clone that hybridizes with the nucleic acid fragment of Claim 1;

(c) isolating the DNA clone identified in step (b); and

(d) sequencing the cDNA or genomic fragment that comprises the clone isolated in step (c)

wherein the sequenced nucleic acid fragment encodes all or a substantial portion of the amino acid sequence encoding a 3-dehydroquinate synthase.

9. A method of obtaining a nucleic acid fragment encoding a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase comprising:

(a) synthesizing an oligonucleotide primer corresponding to a portion of the sequence set forth in any of SEQ ID NOs:1, 3, 5 and 7; and

(b) amplifying a cDNA insert present in a cloning vector using the oligonucleotide primer of step (a) and a primer representing sequences of the cloning vector

wherein the amplified nucleic acid fragment encodes a substantial portion of an amino acid sequence encoding a 3-dehydroquinate synthase.

10. The product of the method of Claim 8.

11. The product of the method of Claim 9.

12. A method for evaluating at least one compound for its ability to inhibit the activity of a 3-dehydroquinate synthase, the method comprising the steps of:

(a) transforming a host cell with a chimeric gene comprising a nucleic acid fragment encoding a 3-dehydroquinate synthase, operably linked to suitable regulatory sequences;

(b) growing the transformed host cell under conditions that are suitable for expression of the chimeric gene wherein expression of the chimeric gene results in production of the 3-dehydroquinate synthase encoded by the operably linked nucleic acid fragment in the transformed host cell;

(c) optionally purifying the 3-dehydroquinate synthase expressed by the transformed host cell;

(d) treating the 3-dehydroquinate synthase with a compound to be tested; and

(e) comparing the activity of the 3-dehydroquinate synthase that has been treated with a test compound to the activity of an untreated 3-dehydroquinate synthase,

thereby selecting compounds with potential for inhibitory activity.

# FIGURE 1

```
SEQ ID NO:9  ME------------------------------------------------------------
SEQ ID NO:2  MAASASSLLAAPASSSCGAISPQLPRGAPA----AASVASPSRHSCYL-LRASPSRRHRS
SEQ ID NO:4  MAAAASSSLLAAASSSSRAAAVSARRAPSASPAAAASLPSPSRASCAPPLRASAARTLRS
SEQ ID NO:6  MASTATNFSLSL----CANQQTPIPKPSFFSNNNHLHFNSNNNWA-----WASVSTSRKS
SEQ ID NO:8  ------------------------------------------------------------
             1                                                          60

                                      *+** *******+*+ ***+  *+ ++++++*+++*++
SEQ ID NO:9  RIV-----------------------VTLGERSYPITIASGLFNEPASFLPLKSGEQVML
SEQ ID NO:2  RFVANAAPTMQPPAESR---VSTVVDVDLGDRSYPIYIGAGLLDEPDLLQRHVHGKRVLV
SEQ ID NO:4  RVVASAAPAMQPPPASR---VSTVVDVDLGDRSYPIYIGAGLLDEPDLLQRHVHGKRVLV
SEQ ID NO:6  RICATSSQVMDPSAAKSEPALPTIVEVDLGSRSYPIYIGSGLLNQPDYLQRHVHGKRVLV
SEQ ID NO:8  ------------------------------------------------------------
             61                                                         120

             *** *+***** **  +*+     *+*+******** ** + +*++**+++ +++ +*
SEQ ID NO:9  VTNETLAPLYLDKVRGVLEQAG--VNVDSVILPDGEQYKSLAVLDTVFTALLQKPHGRDT
SEQ ID NO:2  VTNTTVAPLYLDKVTWALTHNNLNVSVESVILPDGEKYKNMDTLMKVFDKAVESRFDRRC
SEQ ID NO:4  VTNTTVAPLYLEKVTWALTHNNPNVSVESVILPDGEKYKDMGTLMKVFDKAVESRLDRRC
SEQ ID NO:6  VTNETVAPLYLDKVVDALTRGNPNVSVESVILPDGEQYKDMDTLMKVFDKAIESRLDRRC
SEQ ID NO:8  ------------------------------------------------------------AR
             121                                                        180
```

# FIGURE 1 (CONTINUED)

```
                 +********+**++**** ++++***+***+*** ++************+** **** ** *
SEQ ID NO:9      TLVALGGGVVGDLTGFAAASYQRGVRFIQVPTTLLSQVDSSVGGKTAVNHPLGKNMIGAF
SEQ ID NO:2      TFVALGGGVIGDMCGFAAAAFLRGVNFIQIPTTLMAQVDSSVGGKTGINHPLGKNLIGAF
SEQ ID NO:4      TFVALGGGVIGDMCGFAAAAFLRGVNFIQIPTTLMAQVDSSVGGKTGINHPLGKNLIGAF
SEQ ID NO:6      TFVALGGGVIGDMCGFAASAFLRGVNFIQIPTTVMAQVDSSVGGKTGINHRLGKNMIGTF
SEQ ID NO:8      AFVALGGGVIGDMCGFAAAAFLRGVNFIQIPTTLMAQVDSSVGGKTGINHPLGKNLIGAF
                 181                                                      240

                 * *++*++*+ +*+*****+****** *** ***+*+*+ **+*+*+*+   *++ +    *
SEQ ID NO:9      YQPASVVVDLDCLKTLPPRELASGLAEVIKYGIILDGAFFNWLEENLDALLRLDGPAMAY
SEQ ID NO:2      YQPQCVLIDTNTLNTLPDRELASGIAEVVKYGLIRDAPFFEWQEKNMPKLLAREPNALAY
SEQ ID NO:4      YHPQCVLIDTETLNTLPDRELASGIAEVVKYGLIRDAPFFEWQEKNMPALLAREPSALAY
SEQ ID NO:6      YQPQCVLIDTDTLNTLPDRELASGLAEVIKYGLIRDAEFFEWQEKNMHLLLARDPSVMAY
SEQ ID NO:8      YQPQCVLIDTETLNTLPDRELASGVAEVVKYGLIRDAPFFEWQEKNMAAILAREPSALTY
                 241                                                      300

                 +*+*+**+*****  * +*+*****+**************+ *** **********+*** +
SEQ ID NO:9      CIRRCCELKAEVVAADERETGLRALLNLGHTFGHAIEAEMGYGNWLHGEAVAAGMVMAAR
SEQ ID NO:2      AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAIETGTGYGAWLHGEAVAAGTVMATD
SEQ ID NO:4      AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAIETGTGYGAWLHGEAVAAGTVMAAD
SEQ ID NO:6      AIKRSCENKAEVVSLDQKESGLRATLNLGHTFGHAIETGVGYGQWLHGEAVAAGTVMAVD
SEQ ID NO:8      AIKRSCENKAEVVAQDEKESGLRATLNLGHTFGHAIETGLGYGEWLHGEAVAAGTVMAAD
                 301                                                      360
```

## FIGURE 1 (CONTINUED)

```
                 +* ******++ +  +*  ++*  *+**   *   ++  +   *++++++++*++*****+++
SEQ ID NO:9      TSERLGQFSSAETQRIITLLKRAGLPVNGPREMSAQAYLPHMLRDKKVLAGEMRLIL-PL
SEQ ID NO:2      MSHRLGWIDDSIRKRVVDILKQAKLPIAPPETMTVEKFKNIMAVDKKVADGLLRLILLKG
SEQ ID NO:4      MSHRLGWIDESIKKRAIDILEKAKLPITPPEAMTVEKFKSIMAVDKKVADGLLRLILLKG
SEQ ID NO:6      MSYRLGWIDDSLVKRVGDILKQAKLPTAPPETVTVDMFKSVMAVDKKVADGLLRLILLKG
SEQ ID NO:8      MSYRLGWIDESIKKRTFDILDQAKLPVTSPKGMTVEKFRNIMAVDKKVADGLLRLILLKG
                361                                                      420

                 ++* +++++ +           *
SEQ ID NO:9      AIGKSEVRSGVSHELVLNAIAD-CQSA
SEQ ID NO:2      PLGCCVFTGDYDGNALDETLHAFCDN.
SEQ ID NO:4      PLGSCVFTGDYCSS------RSTC-R.
SEQ ID NO:6      PLGNCVFTGDYDRKALDNTLRAFCKS.
SEQ ID NO:8      PLGGCVFTGEYDRKALDETLRAFCDN.
                 421                    447
```

SEQUENCE LISTING

<110> E. I. du Pont de Nemours and Company

<120> Chorismate Biosynthesis Enzymes

<130> BB-1159

<140>
<141>

<150> 60/093,611
<151> July 21, 1998

<160> 9

<170> Microsoft Office 97

<210> 1
<211> 1628
<212> DNA
<213> Zea mays

<400> 1
ccgccgcttt ccgagctctc atggcggcgt ccgcgtcctc cctccttgcc gcccccgcat 60
cctcatcctg tggggccatc tccccccagc tccctcgtgg agctccggcc gctgcctccg 120
tcgcttcgcc gtcgcgtcac tcctgctacc tcctccgcgc ttcccccttcg aggcgccatc 180
ggagtcgctt cgtagccaac gccgctccca ccatgcagcc tccggcggag tccagggtct 240
ccacggtagt cgatgtcgac ctcggcgacc gtagctatcc gatctacatc ggcgcaggcc 300
tcctcgacga gccggacctg ctgcagaggc atgttcatgg taagagggtt ctggtggtga 360
ccaacacgac cgtcgcgccg ctttacctgg acaaggtgac atgggcactc acccacaaca 420
acctgaatgt atcagtggaa agcgtgatcc tgcccgacgg tgaaaagtac aaaaatatgg 480
acacgctgat gaaggtgttt gacaaggcag tcgagtcccg ttttgaccgc cggtgcacat 540
ttgtagcact gggtggtggt gtcattgggg acatgtgtgg ttttgcagct gctgcattcc 600
tccggggcgt caatttcata cagataccaa ctactctgat ggcccaggtg gattcatctg 660
ttggtgggaa aaccgggatt aaccacccac taggaaagaa cttgattggg gcattctacc 720
agccacaatg tgttctaatt gacacaaata cacttaacac attgcctgac agggagctag 780
cttcaggcat tgccgaggta gtaaagtatg ggctcataag ggatgcacca ttctttgagt 840
ggcaagagaa gaacatgccg aaattgttag caagagaacc aaatgctttg gcatatgcta 900
tcaagagatc atgtgaaaac aaagctgaag tggtggcaca agatgagaag gaaagtggcc 960
ttcgagcaac actaaacctg ggtcacacat ttggccatgc tattgagact gggacaggat 1020
atggagcatg gctccatggg gaggctgtcg cagctggaac agttatggca actgacatgt 1080
ctcaccgcct ggggtggata gatgactcca tcagaaaacg tgtggttgac atactaaagc 1140
aagccaaact tcccattgca cctcctgaga ccatgaccgt agagaagttt aaaaacatca 1200
tggctgttga caagaaggtt gctgatggtc tgttgagact catccttctg aaaggaccgc 1260
tagggtgctg tgtatttacg ggggattatg acgggaatgc actcgatgaa accctacatg 1320
cattctgcga caactgagac atcccagttt tggacatcac ttctgtatgc tagtactagg 1380
ctatgtatca atcagagaaa atattgtggt gtacattatc ttcactgcct cattggtata 1440
ttttttgtca gaaataatgg tatattgttt taagtgctgg tttgagtttg tagagaacca 1500
tggctaatca actgcggtca gttgttctct agtaataata gtaaagaagc ttcactacat 1560
cccttctgtc tgttactgtc aagtaacaat actcgtagag atttaaacca aaaaaaaaaa 1620
aaaaaaaa 1628

<210> 2
<211> 438
<212> PRT
<213> Zea mays

<400> 2

```
Met Ala Ala Ser Ala Ser Ser Leu Leu Ala Ala Pro Ala Ser Ser Ser
 1               5                  10                  15

Cys Gly Ala Ile Ser Pro Gln Leu Pro Arg Gly Ala Pro Ala Ala Ala
            20                  25                  30

Ser Val Ala Ser Pro Ser Arg His Ser Cys Tyr Leu Leu Arg Ala Ser
        35                  40                  45

Pro Ser Arg Arg His Arg Ser Arg Phe Val Ala Asn Ala Ala Pro Thr
    50                  55                  60

Met Gln Pro Pro Ala Glu Ser Arg Val Ser Thr Val Val Asp Val Asp
65                  70                  75                  80

Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly Ala Gly Leu Leu Asp
                85                  90                  95

Glu Pro Asp Leu Leu Gln Arg His Val His Gly Lys Arg Val Leu Val
            100                 105                 110

Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Thr Trp
        115                 120                 125

Ala Leu Thr His Asn Asn Leu Asn Val Ser Val Glu Ser Val Ile Leu
    130                 135                 140

Pro Asp Gly Glu Lys Tyr Lys Asn Met Asp Thr Leu Met Lys Val Phe
145                 150                 155                 160

Asp Lys Ala Val Glu Ser Arg Phe Asp Arg Arg Cys Thr Phe Val Ala
                165                 170                 175

Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ala Ala
            180                 185                 190

Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Leu Met Ala
        195                 200                 205

Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Pro Leu
    210                 215                 220

Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln Pro Gln Cys Val Leu Ile
225                 230                 235                 240

Asp Thr Asn Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly
                245                 250                 255

Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg Asp Ala Pro Phe Phe
            260                 265                 270

Glu Trp Gln Glu Lys Asn Met Pro Lys Leu Leu Ala Arg Glu Pro Asn
        275                 280                 285

Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val
    290                 295                 300

Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu
305                 310                 315                 320
```

Gly His Thr Phe Gly His Ala Ile Glu Thr Gly Thr Gly Tyr Gly Ala
325 330 335

Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Thr Asp
340 345 350

Met Ser His Arg Leu Gly Trp Ile Asp Asp Ser Ile Arg Lys Arg Val
355 360 365

Val Asp Ile Leu Lys Gln Ala Lys Leu Pro Ile Ala Pro Pro Glu Thr
370 375 380

Met Thr Val Glu Lys Phe Lys Asn Ile Met Ala Val Asp Lys Lys Val
385 390 395 400

Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Cys
405 410 415

Cys Val Phe Thr Gly Asp Tyr Asp Gly Asn Ala Leu Asp Glu Thr Leu
420 425 430

His Ala Phe Cys Asp Asn
435

<210> 3
<211> 1358
<212> DNA
<213> Oryza sativa

<400> 3
gcacgagtac agcgtcgtct tcgtttcgca gctccaccac ctccatggcg gccgccgcct 60
cctcctctct gctcgccgcc gcctcgtcgt cctcccgcgc ggcggccgtg tccgcccgcc 120
gcgcccccag cgcgagtccg gccgcggcgg cctccctccc ctcgccgtct cgcgcctcct 180
gcgctcctcc cctccgcgct tccgcagcga ggaccctccg cagccgtgtc gtcgcgagcg 240
cggcccccgc tatgcagccg ccgccggcgt cgagggtctc cacggttgtc gacgtcgacc 300
tcggcgaccg gagctacccg atctacatcg gcgcaggcct tctcgacgag cctgacctgc 360
tgcagaggca tgttcatggg aagagggttt tggtggtgac caacaccacc gtcgcgccgc 420
tctacctgga gaaggtgacc tgggcactca cgcacaacaa cccgaatgtt tctgtggaga 480
gcgtgatcct gcccgacggc gagaagtaca aggacatggg cacactgatg aaggttttcg 540
acaaggcagt cgagtcccgg ctggaccggc ggtgcacgtt tgttgcgttg ggaggtggcg 600
ttattgggga catgtgcggt tttgcagctg ctgcgttcct gcgtggtgtc aatttcatac 660
agattcctac tactctgatg gcccaggtgg attcatctgt tggagggaag accggcatta 720
accatccatt ggggaagaac ttaattgggg cgttctacca cccacagtgt gtactgatag 780
acaccgagac actgaataca ttgcctgaca gggaactggc ttcaggcata gctgaggtgg 840
tgaagtatgg tctcataaga gatgcaccgt tctttgaatg gcaagagaaa aacatgccag 900
cattattagc aagagaacca agtgctctgg cctatgctat taagagatcg tgtgaaaaca 960
aagctgaagt ggttgctcag gacgagaagg aaagtggtct ccgagcaaca cttaatctcg 1020
gccatacatt tggccatgct atagaaacag gaactggcta tggagcatgg ctccatgggg 1080
aggctgttgc agctggaaca gttatggcag ctgacatgtc tcaccgcctg ggttggatag 1140
acgagtcaat caagaaacgg gcaattgaca tactagagaa agcgaagctt ccaattacac 1200
ctccagaggc catgacagtg gagaagttca aaagtattat ggccgttgat aagaaggttg 1260
ctgatggatt gctgaggctc atcctcctga aaggacctct gggaagctgt gttttcactg 1320
gcgattactg ttcttcgagg tcgacgtgtc gataaatt 1358

<210> 4
<211> 436
<212> PRT
<213> Oryza sativa

<400> 4

```
Met Ala Ala Ala Ala Ser Ser Ser Leu Leu Ala Ala Ala Ser Ser Ser
 1               5                  10                  15

Ser Arg Ala Ala Ala Val Ser Ala Arg Arg Ala Pro Ser Ala Ser Pro
            20                  25                  30

Ala Ala Ala Ala Ser Leu Pro Ser Pro Ser Arg Ala Ser Cys Ala Pro
        35                  40                  45

Pro Leu Arg Ala Ser Ala Ala Arg Thr Leu Arg Ser Arg Val Val Ala
    50                  55                  60

Ser Ala Ala Pro Ala Met Gln Pro Pro Pro Ala Ser Arg Val Ser Thr
65                  70                  75                  80

Val Val Asp Val Asp Leu Gly Asp Arg Ser Tyr Pro Ile Tyr Ile Gly
                85                  90                  95

Ala Gly Leu Leu Asp Glu Pro Asp Leu Leu Gln Arg His Val His Gly
            100                 105                 110

Lys Arg Val Leu Val Val Thr Asn Thr Thr Val Ala Pro Leu Tyr Leu
        115                 120                 125

Glu Lys Val Thr Trp Ala Leu Thr His Asn Asn Pro Asn Val Ser Val
    130                 135                 140

Glu Ser Val Ile Leu Pro Asp Gly Glu Lys Tyr Lys Asp Met Gly Thr
145                 150                 155                 160

Leu Met Lys Val Phe Asp Lys Ala Val Glu Ser Arg Leu Asp Arg Arg
                165                 170                 175

Cys Thr Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys Gly
            180                 185                 190

Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile Pro
        195                 200                 205

Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr Gly
    210                 215                 220

Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr His Pro
225                 230                 235                 240

Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp Arg
                245                 250                 255

Glu Leu Ala Ser Gly Ile Ala Glu Val Val Lys Tyr Gly Leu Ile Arg
            260                 265                 270

Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Pro Ala Leu Leu
        275                 280                 285

Ala Arg Glu Pro Ser Ala Leu Ala Tyr Ala Ile Lys Arg Ser Cys Glu
    290                 295                 300

Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu Arg
305                 310                 315                 320
```

```
Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr Gly
                325                 330                 335

Thr Gly Tyr Gly Ala Trp Leu His Gly Glu Ala Val Ala Ala Gly Thr
            340                 345                 350

Val Met Ala Ala Asp Met Ser His Arg Leu Gly Trp Ile Asp Glu Ser
        355                 360                 365

Ile Lys Lys Arg Ala Ile Asp Ile Leu Glu Lys Ala Lys Leu Pro Ile
    370                 375                 380

Thr Pro Pro Glu Ala Met Thr Val Glu Lys Phe Lys Ser Ile Met Ala
385                 390                 395                 400

Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys
                405                 410                 415

Gly Pro Leu Gly Ser Cys Val Phe Thr Gly Asp Tyr Cys Ser Ser Arg
            420                 425                 430

Ser Thr Cys Arg
        435
```

<210> 5
<211> 1643
<212> DNA
<213> Glycine max

<400> 5

```
gcacgagccc ggagtggtgt ttgtgttcaa ggaagacagg cgcaggcaca ggcacatcat   60
gttattgtta ttgttggctt cttcactcag caaccaccaa accacaccgt cgcacccatt  120
ccatcgctga ttcccccaat ggcttccact gccaccaatt tctctctttc tctctgcgcc  180
aaccaacaaa ctccaatccc caaaccctct ttcttctcca ataacaacca tttgcacttc  240
aactctaata ataattgggc ctgggcctct gtttccacct ctcgcaagtc aaggatatgc  300
gccacctcct ctcaagttat ggatccctcc gcagcaaaat ccgaacccgc tcttcccacc  360
atcgtcgaag tcgatttggg tagccggagc tatcctattt acatcggatc cgggttacta  420
aaccaacccg actatctcca gaggcatgtg catggaaaga gagtcctagt tgtaactaac  480
gaaaccgttg cgccacttta tctagacaag gttgttgatg ctttgacaag gggaaacccg  540
aatgtttctg tggagagtgt aattttacct gatggtgagc agtacaagga catggatact  600
cttatgaaag tctttgacaa ggccatcgag tcgcggctcg accggcggtg tacatttgtt  660
gctcttggtg gtggtgtgat tggcgacatg tgtggctttg ctgcctctgc cttcctacgt  720
ggtgttaatt ttattcagat tcctacaact gtgatggcac aggtcgattc ttcagttggt  780
ggaaaaactg ggataaatca ccgccttggg aagaacatga tcggtacctt ttaccaacct  840
cagtgtgtgc ttatagacac agacacatta aatacgctac cggataggga actggcatca  900
gggctagcag aggttataaa gtatgggctc attagggatg cagagttttt tgagtggcaa  960
gagaaaaata tgcacttatt attggcaaga gatcctagtg taatggcata tgctataaag 1020
cgatcttgtg aaaacaaggc tgaggttgtg tccttagatc agaaggaaag tggactgagg 1080
gcaacattga acttgggtca tacatttggt catgcaatag aaactggggt aggctatggg 1140
cagtggcttc atggagaggc tgttgcagct ggcacggtaa tggctgttga catgtcatat 1200
cgcctaggtt ggattgatga ttctcttgtg aaacgagttg gagacatttt aaaacaggct 1260
aagttaccca cagcccctcc tgagaccgtg actgtggaca tgtttaaatc tgtcatggca 1320
gtggataaga aggtagcaga tgggttgcta aggcttatcc ttctaaaggg tcctctaggc 1380
aattgtgttt tcacagggga ttatgacaga aaggctctag acaatacgct tcgtgcattc 1440
tgtaaatcct gatctttgtg ctccatgttt ctattgcagc tttatccaat tccttttagt 1500
gagtttttgt atgttgttat ataacctcac ttgtattttt cttatgtact ttggatggca 1560
accatggttc ccattatata cgcaagagta tattatttga aattgaggca tgatcaataa 1620
ttattaaaaa aaaaaaaaaa aaa                                        1643
```

<210> 6
<211> 437
<212> PRT
<213> Glycine max

<400> 6

```
Met Ala Ser Thr Ala Thr Asn Phe Ser Leu Ser Leu Cys Ala Asn Gln
  1               5                  10                  15

Gln Thr Pro Ile Pro Lys Pro Ser Phe Phe Ser Asn Asn Asn His Leu
             20                  25                  30

His Phe Asn Ser Asn Asn Asn Trp Ala Trp Ala Ser Val Ser Thr Ser
         35                  40                  45

Arg Lys Ser Arg Ile Cys Ala Thr Ser Ser Gln Val Met Asp Pro Ser
     50                  55                  60

Ala Ala Lys Ser Glu Pro Ala Leu Pro Thr Ile Val Glu Val Asp Leu
 65                  70                  75                  80

Gly Ser Arg Ser Tyr Pro Ile Tyr Ile Gly Ser Gly Leu Leu Asn Gln
                 85                  90                  95

Pro Asp Tyr Leu Gln Arg His Val His Gly Lys Arg Val Leu Val Val
            100                 105                 110

Thr Asn Glu Thr Val Ala Pro Leu Tyr Leu Asp Lys Val Val Asp Ala
        115                 120                 125

Leu Thr Arg Gly Asn Pro Asn Val Ser Val Glu Ser Val Ile Leu Pro
    130                 135                 140

Asp Gly Glu Gln Tyr Lys Asp Met Asp Thr Leu Met Lys Val Phe Asp
145                 150                 155                 160

Lys Ala Ile Glu Ser Arg Leu Asp Arg Arg Cys Thr Phe Val Ala Leu
                165                 170                 175

Gly Gly Gly Val Ile Gly Asp Met Cys Gly Phe Ala Ala Ser Ala Phe
            180                 185                 190

Leu Arg Gly Val Asn Phe Ile Gln Ile Pro Thr Thr Val Met Ala Gln
        195                 200                 205

Val Asp Ser Ser Val Gly Gly Lys Thr Gly Ile Asn His Arg Leu Gly
    210                 215                 220

Lys Asn Met Ile Gly Thr Phe Tyr Gln Pro Gln Cys Val Leu Ile Asp
225                 230                 235                 240

Thr Asp Thr Leu Asn Thr Leu Pro Asp Arg Glu Leu Ala Ser Gly Leu
                245                 250                 255

Ala Glu Val Ile Lys Tyr Gly Leu Ile Arg Asp Ala Glu Phe Phe Glu
            260                 265                 270

Trp Gln Glu Lys Asn Met His Leu Leu Leu Ala Arg Asp Pro Ser Val
        275                 280                 285
```

```
Met Ala Tyr Ala Ile Lys Arg Ser Cys Glu Asn Lys Ala Glu Val Val
    290                 295                 300

Ser Leu Asp Gln Lys Glu Ser Gly Leu Arg Ala Thr Leu Asn Leu Gly
305                 310                 315                 320

His Thr Phe Gly His Ala Ile Glu Thr Gly Val Gly Tyr Gly Gln Trp
                325                 330                 335

Leu His Gly Glu Ala Val Ala Ala Gly Thr Val Met Ala Val Asp Met
            340                 345                 350

Ser Tyr Arg Leu Gly Trp Ile Asp Asp Ser Leu Val Lys Arg Val Gly
        355                 360                 365

Asp Ile Leu Lys Gln Ala Lys Leu Pro Thr Ala Pro Pro Glu Thr Val
    370                 375                 380

Thr Val Asp Met Phe Lys Ser Val Met Ala Val Asp Lys Lys Val Ala
385                 390                 395                 400

Asp Gly Leu Leu Arg Leu Ile Leu Leu Lys Gly Pro Leu Gly Asn Cys
                405                 410                 415

Val Phe Thr Gly Asp Tyr Asp Arg Lys Ala Leu Asp Asn Thr Leu Arg
            420                 425                 430

Ala Phe Cys Lys Ser
        435
```

<210> 7
<211> 1103
<212> DNA
<213> Triticum aestivum

<400> 7

```
gcacgagcat tcgttgcgct aggtggcggt gtcattgggg acatgtgtgg gtttgcagct   60
gctgcattcc tccgtggtgt caatttcata caaataccca caactctgat ggcccaggtg  120
gattcttctg ttggagggaa gactgggatt aaccatccac tgggaaagaa cttaattggc  180
gcattctacc agccgcagtg tgtactcatt gacacagaga cactgaatac attgcctgac  240
agagagctgg cttcaggtgt cgctgaggtg gtgaagtatg gtctcataag agatgcaccc  300
ttctttgagt ggcaagaaaa gaatatggca gcaatcttag cgagagaacc aagtgccctg  360
acctacgcta taaagagatc atgtgaaaac aaagctgaag ttgttgctca ggatgagaag  420
gaaagtggtc ttcgagcaac actaaacctg ggtcatacat ttggtcatgc tatagaaaca  480
gggcttggct atggagaatg gctccacggg gaggctgttg ctgctggaac ggttatggca  540
gctgacatgt cttaccgcct gggctggata gacgagtcca tcaagaaacg gacatttgac  600
atactagatc aagccaagct tcccgtgaca tcaccaaagg gcatgacggt agagaagttc  660
agaaacatca tggcggtcga caagaaggtc gcggacggat tgttgaggct catccttctg  720
aaaggacctc tgggaggctg cgttttcacc ggcgagtacg acaggaaagc cctggacgaa  780
accctccgcg ctttctgcga caactgatcg ttgtcaagct ggagctccag catttctgta  840
tactagtagg ctgtaatatc atatcagaga gagagagaga gagagtgcat tgtgtgctac  900
ctttattttt tcatttcgtt ttggtgagaa taatggaatg cggtaagggg ttggttccag  960
gtcgtagtac taccaattca ttgtagggca tcatgcctca gtaatcgtga atcttgatca 1020
gtcattttga gaaacacaat tcggaacgag atcctctgac gtaggaagtt ttaaaaaaag 1080
agaagaaaaa aaaaaaaaaa aaa                                        1103
```

<210> 8
<211> 268
<212> PRT
<213> Triticum aestivum

<400> 8

```
Ala Arg Ala Phe Val Ala Leu Gly Gly Gly Val Ile Gly Asp Met Cys
 1               5                  10                  15

Gly Phe Ala Ala Ala Ala Phe Leu Arg Gly Val Asn Phe Ile Gln Ile
            20                  25                  30

Pro Thr Thr Leu Met Ala Gln Val Asp Ser Ser Val Gly Gly Lys Thr
        35                  40                  45

Gly Ile Asn His Pro Leu Gly Lys Asn Leu Ile Gly Ala Phe Tyr Gln
    50                  55                  60

Pro Gln Cys Val Leu Ile Asp Thr Glu Thr Leu Asn Thr Leu Pro Asp
65                  70                  75                  80

Arg Glu Leu Ala Ser Gly Val Ala Glu Val Val Lys Tyr Gly Leu Ile
                85                  90                  95

Arg Asp Ala Pro Phe Phe Glu Trp Gln Glu Lys Asn Met Ala Ala Ile
            100                 105                 110

Leu Ala Arg Glu Pro Ser Ala Leu Thr Tyr Ala Ile Lys Arg Ser Cys
        115                 120                 125

Glu Asn Lys Ala Glu Val Val Ala Gln Asp Glu Lys Glu Ser Gly Leu
    130                 135                 140

Arg Ala Thr Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Thr
145                 150                 155                 160

Gly Leu Gly Tyr Gly Glu Trp Leu His Gly Glu Ala Val Ala Ala Gly
                165                 170                 175

Thr Val Met Ala Ala Asp Met Ser Tyr Arg Leu Gly Trp Ile Asp Glu
            180                 185                 190

Ser Ile Lys Lys Arg Thr Phe Asp Ile Leu Asp Gln Ala Lys Leu Pro
        195                 200                 205

Val Thr Ser Pro Lys Gly Met Thr Val Glu Lys Phe Arg Asn Ile Met
    210                 215                 220

Ala Val Asp Lys Lys Val Ala Asp Gly Leu Leu Arg Leu Ile Leu Leu
225                 230                 235                 240

Lys Gly Pro Leu Gly Gly Cys Val Phe Thr Gly Glu Tyr Asp Arg Lys
                245                 250                 255

Ala Leu Asp Glu Thr Leu Arg Ala Phe Cys Asp Asn
            260                 265
```

<210> 9
<211> 362
<212> PRT
<213> Escherichia coli

<400> 9

```
Met Glu Arg Ile Val Val Thr Leu Gly Glu Arg Ser Tyr Pro Ile Thr
 1               5                  10                  15

Ile Ala Ser Gly Leu Phe Asn Glu Pro Ala Ser Phe Leu Pro Leu Lys
            20                  25                  30

Ser Gly Glu Gln Val Met Leu Val Thr Asn Glu Thr Leu Ala Pro Leu
        35                  40                  45

Tyr Leu Asp Lys Val Arg Gly Val Leu Glu Gln Ala Gly Val Asn Val
    50                  55                  60

Asp Ser Val Ile Leu Pro Asp Gly Glu Gln Tyr Lys Ser Leu Ala Val
65                  70                  75                  80

Leu Asp Thr Val Phe Thr Ala Leu Leu Gln Lys Pro His Gly Arg Asp
                85                  90                  95

Thr Thr Leu Val Ala Leu Gly Gly Gly Val Val Gly Asp Leu Thr Gly
            100                 105                 110

Phe Ala Ala Ala Ser Tyr Gln Arg Gly Val Arg Phe Ile Gln Val Pro
        115                 120                 125

Thr Thr Leu Leu Ser Gln Val Asp Ser Ser Val Gly Gly Lys Thr Ala
    130                 135                 140

Val Asn His Pro Leu Gly Lys Asn Met Ile Gly Ala Phe Tyr Gln Pro
145                 150                 155                 160

Ala Ser Val Val Val Asp Leu Asp Cys Leu Lys Thr Leu Pro Pro Arg
                165                 170                 175

Glu Leu Ala Ser Gly Leu Ala Glu Val Ile Lys Tyr Gly Ile Ile Leu
            180                 185                 190

Asp Gly Ala Phe Phe Asn Trp Leu Glu Glu Asn Leu Asp Ala Leu Leu
        195                 200                 205

Arg Leu Asp Gly Pro Ala Met Ala Tyr Cys Ile Arg Arg Cys Cys Glu
    210                 215                 220

Leu Lys Ala Glu Val Val Ala Ala Asp Glu Arg Glu Thr Gly Leu Arg
225                 230                 235                 240

Ala Leu Leu Asn Leu Gly His Thr Phe Gly His Ala Ile Glu Ala Glu
                245                 250                 255

Met Gly Tyr Gly Asn Trp Leu His Gly Glu Ala Val Ala Ala Gly Met
            260                 265                 270

Val Met Ala Ala Arg Thr Ser Glu Arg Leu Gly Gln Phe Ser Ser Ala
        275                 280                 285

Glu Thr Gln Arg Ile Ile Thr Leu Leu Lys Arg Ala Gly Leu Pro Val
    290                 295                 300

Asn Gly Pro Arg Glu Met Ser Ala Gln Ala Tyr Leu Pro His Met Leu
305                 310                 315                 320
```

Arg Asp Lys Lys Val Leu Ala Gly Glu Met Arg Leu Ile Leu Pro Leu
325 330 335

Ala Ile Gly Lys Ser Glu Val Arg Ser Gly Val Ser His Glu Leu Val
340 345 350

Leu Asn Ala Ile Ala Asp Cys Gln Ser Ala
355 360

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

# PCT

## NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

Assistant Commissioner for Patents
United States Patent and Trademark Office
Box PCT
Washington, D.C.20231
ETATS-UNIS D'AMERIQUE

in its capacity as elected Office

| | |
|---|---|
| **Date of mailing** (day/month/year)<br>21 March 2000 (21.03.00) | |
| **International application No.**<br>PCT/US99/16354 | **Applicant's or agent's file reference**<br>BB1159 |
| **International filing date** (day/month/year)<br>20 July 1999 (20.07.99) | **Priority date** (day/month/year)<br>21 July 1998 (21.07.98) |
| **Applicant**<br>CAHOON, Rebecca, E. et al | |

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

07 February 2000 (07.02.00)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was

[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

**The International Bureau of WIPO**
**34, chemin des Colombettes**
**1211 Geneva 20, Switzerland**

Facsimile No.: (41-22) 740.14.35

Authorized officer

Dominique DELMAS

Telephone No.: (41-22) 338.83.38